**UNIVERSITY OF SOUTHAMPTON**

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING

Electronics and Computer Science

# Automorphisms of Random Recursive Trees

by

**David Matthews**

Thesis for the degree of Doctor of Philosophy

April 2017

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING
Electronics and Computer Science

Doctor of Philosophy

AUTOMORPHISMS OF RANDOM RECURSIVE TREES

by David Matthews

A *Viswanath random Fibonacci sequence* Viswanath random Fibonacci sequence is defined by $f_0 = f_1 = 1$ and $f_n = \pm f_{n-1} + f_{n-2}$ for $n > 1$, where addition and subtraction are chosen by flipping a fair coin. Viswanath showed

$$|f_n|^{\frac{1}{n}} \xrightarrow{\text{a.s.}} 1.13198824\ldots$$

The number $\mathcal{V} = 1.13198824\ldots$ has subsequently become known as Viswanath's constant.

A *random recursive tree* is a nested family of rooted trees $\{t_n\}_{n\in\mathbb{N}}$ where each $t_n$ is a rooted tree on $n$ vertices and $t_{n+1}$ is obtained from $t_n$ by adding a new vertex via an edge to a vertex of $t_n$ chosen uniformly at random.

The symmetry of a rooted tree $t$ is encoded by $\mathrm{Aut}(t)$, the automorphism group of $t$. MacArthur and Anderson identified what they call "an intriguing relationship" between random Fibonacci sequences and the automorphism group of random recursive trees that we investigate in this thesis.

# Contents


**Declaration of Authorship** **vii**

**Acknowledgements** **ix**

**1 Introduction** **1**

1.1 Random Fibonacci sequences . . . 1
1.2 Random graph theory . . . 3
1.2.1 Automorphisms of random graphs . . . 4
1.2.2 Random recursive trees . . . 5
1.3 Random recursive trees and random Fibonacci sequences . . . 9
1.4 Results . . . 10

**2 Background** **13**

2.1 The automorphism group of a rooted forest . . . 14
2.1.1 The rooted geometric decomposition . . . 15
2.1.2 The elementary and non-elementary subgroups . . . 17
2.2 Rooted ordered trees . . . 17
2.2.1 Historical orderings . . . 19
2.2.2 Weight function . . . 20
2.3 Two models of randomly growing trees . . . 20
2.3.1 A restriction on the weight function . . . 22

**3 Two models of random attachment trees** **25**

3.1 Correspondence between saplings and historical sequences . . . 26
3.2 Network motifs . . . 30

**4 Existence of limit** **35**

4.1 Branching processes . . . 35
4.2 An application of Theorem 2.3.3 . . . 38

**5 Automorphisms of random recursive trees** **43**

5.1 Limiting contribution of $k$-stars . . . 44
5.2 A correspondence between $(n+1)$-saplings and $S_n$ . . . 45
5.3 The cycle indicator polynomial . . . 49
5.3.1 A family of generating functions . . . 52
5.4 Recursively exponential polynomials . . . 54
5.5 Proof of Theorem 3 . . . 60

**6 Expectation** **67**
6.1 A generating function approach . . . . . . . . . . . . . . . . . . . . . . . . 68
6.2 Differential posets . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 71

**References** **75**

## Declaration of Authorship

I, David Matthews , declare that the thesis entitled *Automorphisms of Random Recursive Trees* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission

Signed:.......................................................................................................................

Date:..........................................................................................................................

## Acknowledgements

Firstly and most importantly, thanks to my supervisor James “Jim” Anderson for our wide-ranging conversations which introduced me to a kaleidoscope of fascinating mathematics. Without you this thesis would not have been possible. I would also like to thank my adviser Ben MacArthur whose conjectures I have lived and breathed for the last four years.

Thanks to everyone I have met that inhabits or has inhabited 54/7035 and 54/8037 for your great curiosity and companionship. In particular Thomas Kenneth Marvolo Harris (K-theory enthusiast, writer of letters and housemate of many glorious years) Alex Bailey (for your puzzles and for introducing me to Perudo), Joe J. Tait (who loves rings), Charles Cox (sharer of mathematics), Mike Winston Thermidor West (for his explanations and family values), Hector Durham (for whom group theory is always the hammer whatever the nail), Martin Finn-Sell (keen on graphs and burgers) and Matt Burfitt (Perudo world champion).

Finally I would like to thank Lotta Viika for enduring (amongst a myriad of other things) my ramblings on the subject of trees and random Fibonacci sequences.

# Chapter 1

# Introduction

MacArthur and Anderson identify what they call "an intriguing relationship" in [MA06] between *random Fibonacci sequences* and *random recursive trees*. This relationship provides the motivation for this thesis. In Section 1.1 we define the random Fibonacci sequence and we go on to define random recursive trees in Section 1.2. In Section 1.3 we bring these together and state MacArthur and Anderson's findings. We end this Chapter with Section 1.4 in which we state our main results and give a brief overview of this thesis.

**Notation.** Throughout this document we write

$$\mathbb{R}^{>} := \{x \in \mathbb{R} : x > 0\}$$
$$\mathbb{R}^{\geq} := \{x \in \mathbb{R} : x \geq 0\}$$

For clarity we will write $\mathbb{N}$ for the positive integers and define $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Define $[n] := \{1, 2, \ldots, n\}$.

We adopt the standard notation, $S_n$, for the symmetric group on $n$ elements .

Suppose that $\{X_n\}$ is a sequence of real-valued random variables and $X \in \mathbb{R}$. We write $X_n \xrightarrow{\text{a.s.}} X$ if $X_n$ converges to $X$ *almost surely* [Bil13].

## 1.1 Random Fibonacci sequences

The *Fibonacci sequence* is probably the most famous sequence in mathematics. It is defined by $g_0 = g_1 := 1$ and $g_n := g_{n-1} + g_{n-2}$ for $n \geq 2$. Each element of Fibonacci's sequence is known as a *Fibonacci number*.

Recently a variant of Fibonacci's sequence called the *random Fibonacci sequences* have excited a great deal of interest [Vis00, SK01, ET99]. A random Fibonacci sequence is defined by $f_0 = f_1 := 1$ and subsequently by the recursion

$$f_n := \pm\alpha f_{n-1} \pm \beta f_{n-2}, \tag{1.1.1}$$

where $\alpha, \beta \in \mathbb{R}$ and addition and subtraction are chosen according to some probability distribution for $n \geq 2$.

In [Vis00], Viswanath investigated the behaviour of the absolute value $|f_n|$ with $\alpha = \beta = 1$ and where each plus and minus is chosen independently with probability $\frac{1}{2}$. Since $|f_n|$ is either $|f_{n-1}| + |f_{n-2}|$ or $||f_{n-1}| - |f_{n-2}||$ it is enough to consider the recurrence

$$f_n = \pm f_{n-1} + f_{n-2},$$

where addition and subtraction are chosen with probability $\frac{1}{2}$. We call these sequences the *Viswanath random Fibonacci sequences*. We may rewrite a Viswanath random Fibonacci sequence in terms of a random matrix product as follows:

$$\begin{pmatrix} f_{n-1} \\ f_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} \begin{pmatrix} f_{n-2} \\ f_{n-1} \end{pmatrix}.$$

More rigorously we define two matrices,

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}.$$

Let $\mu$ denote the probability distribution that chooses $A$ or $B$ with probability $\frac{1}{2}$. Then,

$$\begin{pmatrix} f_{n-1} \\ f_n \end{pmatrix} = M_{n-1}M_{n-2}\ldots M_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

where $M_{n-1}M_{n-2}\ldots M_1$ is a product of independent identically distributed random matrices; each chosen from $\{A, B\}$ with distribution $\mu$. The problem of calculating the limiting value of $|f_n|^{\frac{1}{n}}$ can thus be written in terms of the random matrix product $M_nM_{n-1}\ldots M_1$.

One consequence of the Fürstenberg-Kesten Theorem [FK60] is that given such a pair of matrices and a measure $\mu$, there exists a constant $\lambda$ such that,

$$\frac{1}{n}\log ||M_nM_{n-1}\ldots M_1|| \xrightarrow{\text{a.s.}} \lambda$$

where $||M_i||$ is the usual 2-norm of a matrix (see [Mey00]). The constant $\lambda$ is called a *Lyapunov Exponent*. The interested reader should see [KR68] for a definition and discussion of Lyapunov exponents.

Viswanath used machinery from the theory of random matrix products developed by

Furstenberg and Kesten in [FK60] together with an ingenious application of a Stern-Brocot tree (defined in [GPK94]) to prove,

$$|f_n|^{\frac{1}{n}} \xrightarrow{\text{a.s.}} 1.13198824\ldots$$

The number $\mathcal{V} = 1.13198824\ldots$ has subsequently become known as *Viswanath's constant* [ET99]. In [Vis00] Viswanath's constant is calculated by evaluating an integral. No closed form solution of this integral is known hence, more precisely, Viswanath proves that $\mathcal{V}$ is in the interval $[1.13198824, 1.13198825)$.

In [ET99], Embree and Trefethen generalise the work done by Viswanath. They investigate the 1-parameter family of random Fibonacci sequences given by the recurrence,

$$f_n = f_{n-1} \pm \beta f_{n-2},$$

for $\beta \in \mathbb{R}$ and where addition and subtraction are both chosen with probability $\frac{1}{2}$. Embree and Trefethen [ET99], conjectured the following:

(i) If $0 < \beta < 0.702582$ then there exists some constant $C_\beta \in \mathbb{R}$ such that

$$|f_n|^{\frac{1}{n}} \xrightarrow{\text{a.s.}} C_\beta.$$

(ii) If $0.702585 < \beta < \infty$ then there exists some constant $C_\beta \in \mathbb{R}$ such that

$$|f_n|^{\frac{1}{n}} \xrightarrow{\text{a.s.}} C_\beta.$$

(iii) There exists a threshold value $\beta^* \in (0.702582, 0.702585)$ such that the corresponding random Fibonacci sequence neither grows exponentially nor decays exponentially.

Besides random matrix products random Fibonacci sequences appear in many areas of mathematics. For example, we can think of $A$ and $B$ as Möbius transformations of the complex plane. In this case $f_n$ corresponds to a random composition of Möbius transformations. In addition we can think of $A$ and $B$ acting on the integer lattice [KN10]. In this case the random matrix product $f_n$ corresponds to a random walk on the integer lattice.

## 1.2 Random graph theory

The theory of random graphs began in 1959 with Erdős and Rényi's seminal paper [ER59]. Between the 1960s and the 1990s the majority of graph theory focused on *Erdős-Rényi random graphs*, characterised by a binomial degree distribution, and *completely* *regular graphs* in which all vertices have the same degree. The invention and subsequent prevalence of the World Wide Web in the 1990s together with greater computational power contributed to a renewed interest in the study of random graphs under the guise of *network science* [WS98, AB00, BAJ00, AB02].

Graphs have been devised as null models for phenomena as diverse as the Internet [AJB99], the World Wide Web [AH00], actor collaboration [BA99], scientific collaboration [New01], ecological and cellular activity [JBA+06, BO04], scientific citations [LJM+07], power systems [Hol06] and protein interactions [WS98]. Whilst these *real-world* networks are certainly not regular, neither can they be modelled by an Erdős-Rényi random graph [WS98].

Astonishingly, almost all of the networks we mentioned above share three characteristics [BAJ00, WS98]:

(i) low *average path length*,

(ii) high *cluster coefficient* and

(iii) a *degree distribution* that approximately follows an *inverse power law*.

The *Barabási-Albert* and the *Watts and Strogatz* models are defined in [BAJ00] and [WS98] respectively. An interesting refinement of the Barabási-Albert graph model was proposed by Bollobás in [BBCR03].

Barabási-Albert random graphs have a degree distribution that follows an inverse power law and exhibit low average path length [BAJ00]. Watts and Strogatz random graphs have a low average path length and a high cluster coefficient [WS98]. The Barabási-Albert and the Watts and Strogatz models are *dynamic*, i.e. they change over time like many real-world networks. The Barabási-Albert model features *preferential attachment* where as the graph grows new vertices are more likely to be attached to an existing vertex with a high degree than to a vertex with a low degree. This "rich get richer" phenomena is well known in real-world networks [BAJ00].

### 1.2.1 Automorphisms of random graphs

It is a Theorem of Cameron [C+04] that almost all graphs have no non-trivial automorphisms. That is, the proportion of graphs on $n$ vertices which have a non-trivial automorphism tends to zero as $n \to \infty$. In fact this is true for both labelled and unlabelled graphs. On the other hand MacArthur, Sánchez-García and Anderson calculated the order of the automorphism group of many real-world networks and found automorphism groups of a large order (relative to the number of vertices) [MSGA08].

MacArthur, Sánchez-García and Anderson suggested that real-world networks that grow by the addition of vertices and edges under preferential attachment are naturally tree-like in [MSGA08]. If real-world networks are indeed tree-like then one could understand properties of real-world networks by understanding properties of random trees.

Random characteristics of graphs such as degree distribution, clustering coefficient and average path length have been well studied, for example see [BAJ00] and [WS98]. In this thesis we will investigate symmetry (characterised by the automorphism group) of a family of random graph models called random recursive trees.

### 1.2.2 Random recursive trees

The second part of the "intriguing relationship" identified by MacArthur and Anderson are the *random recursive trees*, which we define in this section.

For definitions and description of *rooted graphs*, *rooted trees* and *rooted forests* the reader should see [Bol13]. Definition 1.2.1 (rooted automorphism) is a natural extension of the standard definition of a rooted tree automorphism. Definitions in the latter part of this section including *attachement tree* and *tree multiplicity* are taken from the theory of Hopf algebras of rooted trees [Hof03].

This section is particularly important because we define three functions that are central to this thesis. The first is Definition 1.2.2 (root permutation function) which counts the rooted automorphisms that permute branches around the root vertex. The second key function is Definition 1.2.3 (weight function) which acts as the attachment parameter in the growth of random attachment trees. We take the definition of the weight function directly from [RTV07]. The final function of note, Definition 1.2.8 (tree multipilicity), is a standard combinatorial function used by Connes, Kreimer and others in the study of Hopf algebras of rooted trees [CK00, CK99, Hof03].

A *rooted graph* $G$ is a triple $G = (r, V, E)$ with *vertex set* $V$, *edge set* $E$ and a single specified vertex $r$ called the *root*. If it is not immediately obvious which rooted graph we are talking about we write $V(G), E(G)$ and $r(G)$ for the vertex set, the edge set and the root of $G$ respectively.

Given a pair of vertices $v, w \in V$ we write $(v, w)$ for the edge with termini $v$ and $w$. For succinctness we write $|G| := |V(G)|$ for the number of vertices in any rooted graph $G$ .

A *rooted tree* $t$ is a connected rooted graph that admits no cycles. A rooted forest $f$ is the disjoint union of rooted trees.

Let $\mathcal{R}_n$ denote the set of all rooted trees on $n$ vertices and define,

$$\mathcal{R} := \bigcup_{n\in\mathbb{N}} \mathcal{R}_n.$$

A study of rooted trees as geometric objects necessarily involves the study of their symmetries described by their automorphism group.

**Definition 1.2.1** (rooted automorphism)**.** Let $t_1, t_2, \dots, t_k$ be rooted trees that are pairwise non-isomorphic. Suppose $f$ is the disjoint union of $\alpha_i$ copies of $t_i$ for $i = 1, 2, 3, \dots, k$. Label these disjoint rooted trees $t_{1,1}, t_{1,2}, \dots, t_{1,\alpha_1}, t_{2,1}, \dots, t_{2,\alpha_2}, \dots, t_{k,1}, \dots, t_{k,\alpha_k}$. Each rooted tree $t_{i,j} = (r_{i,j}, V_{i,j}, E_{i,j})$. A map $a : V(f) \to V(f)$ is an *rooted automorphism* of $f$ if and only if

- $a\,(r_{i,j}) = r_{i',j'}$ for some $i, i' \in [k]$ and $j \in [\alpha_i]$ and $j' \in [\alpha_{i'}]$. And
- $(v, w) \in E(f)$ if and only if $(a(v), a(w)) \in E(f)$.

We write $f \cong f'$ if there exists a rooted tree isomorphism $a : V(f) \to V(f')$ and we say that $f$ is *isomorphic* to $f'$.

It is well known that the set of rooted automorphisms (together with binary operation composition of maps) of $f$ form a group [Bol13] called the *automorphism group*, denoted $\mathrm{Aut}(f)$. We denote the order of the automorphism group

$$\zeta(f) = |\mathrm{Aut}(f)|. \tag{1.2.1}$$

More precisely then, $\mathcal{R}_n$ consists of *isomorphism classes* of rooted trees with $n$ vertices. For example $\mathcal{R}_1$ has one element, namely the rooted tree on one vertex and no edges which we denote $\bullet$ . In Figure 1.1 we depict $\mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_3 \cup \mathcal{R}_4$.

Suppose $f$ is the rooted forest described in Definition 1.2.1. Let $t = (r, V, E)$ be the rooted tree with vertices $V(t) = V(f) \cup \{r\}$ and edges

$$E(t) = E(f) \cup \bigcup_{i=1}^{k} \left( \bigcup_{j=1}^{\alpha_i} (r, r_{i,j}) \right).$$

We have built $t$ by taking $f$, adding a new vertex declared to be the root vertex $r(t)$ and joining $r(t)$ via an edge to the root of every connected component of $f$. We write,

$$t := B^+(t_1^{\alpha_1}, t_2^{\alpha_2}, \dots, t_k^{\alpha_k}) \tag{1.2.2}$$

This notation is also used in [CK98] and [Bro04] .

**Definition 1.2.2.** We define the *root permutation function* $\Lambda : \mathcal{R} \to \mathbb{N}$ by

$$\Lambda(t) := \prod_{i=1}^{m} (\alpha_i!) \tag{1.2.3}$$

where $t = B^+(t_1^{\alpha_1}, \dots, t_m^{\alpha_m})$ .

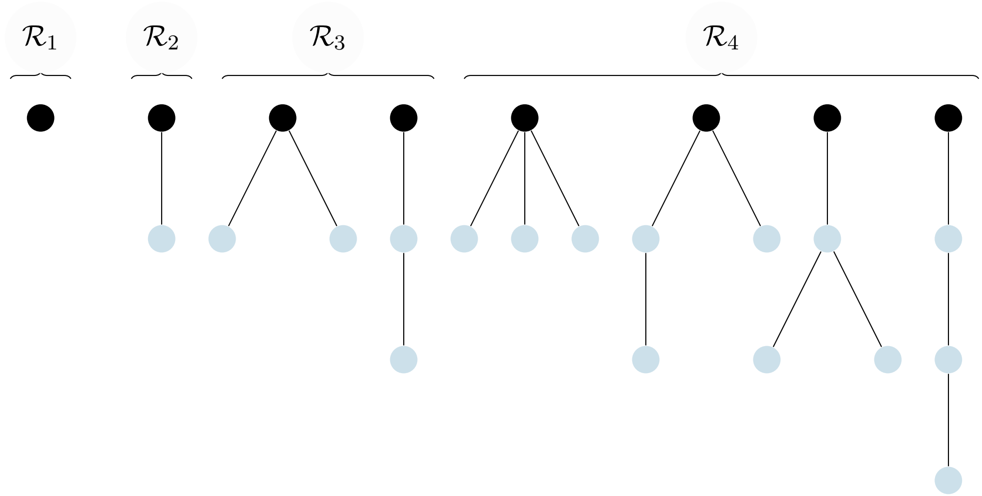


Figure 1.1: $\mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_3 \cup \mathcal{R}_4$. We have indicated the root vertex of each tree by shading it black. As we can see $|\mathcal{R}_1| = 1$, $|\mathcal{R}_2| = 1$, $|\mathcal{R}_3| = 2$ and $|\mathcal{R}_4| = 4$.

Suppose $t = (r, V, E) \in \mathcal{R}$ and $t' = (r', V', E') \in \mathcal{R}$. If $r = r'$ and $V' = V \cup \{w\}$ and $E' = E \cup \{(v, w)\}$ for some $v \in V$ we write $t \lhd t'$. In this case we say that $t'$ can be obtained from $t$ by adding a single edge incident to $v \in V$ and an additional vertex $w$. For example suppose $t$ is the unique rooted tree on 1 vertex and $t'$ is the unique rooted tree on 2 vertices shown in Figure 1.1. Then $t \lhd t'$. This notation is taken from [Hof03].

**Definition 1.2.3** (weight function)**.** A *weight function* $w : \mathbb{N} \to \mathbb{R}$ satisfies $w(i) > 0$ for all $i \in \mathbb{N}$.

For any pair $u, v \in V$ let $d(u, v)$ denote the usual graph distance. There is a natural orientation on every rooted tree. Suppose $t = (r, V, E)$ is a rooted tree. Given an edge $(v, w) \in E$ without loss of generality suppose $d(r, v) + 1 = d(r, w)$. Orient each edge from $v$ to $w$. In other words each edge is oriented away from the root vertex. The *outdegree* $\text{out}(v, t)$ of a vertex $v \in V(t)$ is defined to be

$$\text{out}(v, t) := |\{w \in V : (v, w) \in E \text{ and } d(r, v) + 1 = d(r, w)\}| .$$

The orientation on the graph draws a natural parallel with phylogenetic trees hence it is common to use the following terminology. Suppose $(v, w) \in E$ and $d(r, v) + 1 = d(r, w)$. We say that $v$ is the *parent* of $w$ and $w$ is the *child* of $v$. The set $\text{Ch}_v$ of all vertices $w$ such that $w$ is the child of $v$ is called the *children* of $v$.

**Definition 1.2.4** (attachment tree)**.** Suppose $w : \mathbb{N} \to \mathbb{R}$ is a weight function. An *attachment tree* $\{t_n\}_{n \in \mathbb{N}}$ is a sequence of rooted trees,

$$\bullet = t_1 \lhd t_2 \lhd \cdots \lhd t_n \lhd \dots$$

constructed as follows: At time $n = 1$ begin with the rooted tree $t_1 = \bullet$. At each subsequent time $n \in \mathbb{N}$ a vertex $u_{n-1} \in V(t_{n-1})$ is chosen with probability

$$p = \frac{w(\text{out}(u_{n-1}, t_{n-1}))}{\sum_{v \in V(t_{n-1})} w(\text{out}(v, t_{n-1}))}.$$

The rooted tree $t_n$ is constructed by attaching a new vertex to $v_{n-1}$.

We remark that if $\{t_n\}_{n\in\mathbb{N}}$ is an attachment tree then $|t_n| = n$ for all $n \in \mathbb{N}$. Two attachment trees $\{t_n\}_{n\in\mathbb{N}}$ and $\{t'_n\}_{n\in\mathbb{N}}$ are considered the same and called *isomorphic* if $t_n \cong t'_n$ for all $n \in \mathbb{N}$. We denote the set of all (isomorphism classes of) attachment trees with weight function $w$ by $\mathcal{T}_w$.

**Example 1.2.5.** (i) A *random plane recursive tree* is the attachment tree determined by weight function $w(n) = n+1$ for all $n \in \mathbb{N}$ [KRL00] .We can think of the children of each vertex ordered (clockwise), and new vertices inserted at any place in this ordering. In particular if $\text{out}(v, t_{n-1}) = d$ then there are $d + 1$ possible places to add a new vertex. In total a plane tree with $n$ vertices has $2n - 1$ such places and each of these places is chosen uniformly at random.

(ii) A *random recursive trees* is the attachment tree determined by weight function $\widetilde{w}(n) = 1$ for all $n \in \mathbb{N}$. We can think of each new vertex being attached to a vertex chosen uniformly at random from the existing vertices.

Random recursive trees will be our primary objects of study in this thesis. A random recursive tree is also called a *Yule Tree*, after British mathematician George Udny Yule who investigated it (and other random attachment schemes) much earlier than Barabási and others.

Let $\{t_n\}_{n\in\mathbb{N}}$ be a random recursive tree and $X_{n,i}(\{t_n\}_{n\in\mathbb{N}})$ be the number of vertices with outdegree $i(\geq 0)$ in $t_n$. In [Jan05], Janson used a *Pólya urn model* (see [Mah08]) to prove the following two Theorems.

**Theorem 1.2.6.** [Jan05, Theorem 1.1]. Let $i \in \mathbb{N}_0$ and $\{t_n\}_{n\in\mathbb{N}}$ be a random recursive tree. Then,

$$\frac{X_{n,i}}{n} \xrightarrow{\text{a.s.}} 2^{-(i+1)}.$$

Thus the degree distribution in a random recursive tree follows an inverse power law. Let $Y_{n,i}$ be the number of vertices of outdegree $i \geq 0$ in a random plane recursive tree on $n$ vertices.

**Theorem 1.2.7.** [Jan05, Theorem 1.3]. Let $i \in \mathbb{N}_0$. Then,

$$\frac{Y_{n,i}}{n} \xrightarrow{\text{a.s.}} \frac{4}{(i+1)(i+2)(i+3)}$$

Suppose $\{t_i\}_{i\in\mathbb{N}}$ is an attachment tree and fix $n \in \mathbb{N}$. The *n-sapling* $\{t_i\}_{i=1}^n$ is the nested sequence of rooted trees,

$$\{t_i\}_{i=1}^n = \bullet \lhd t_2 \lhd \cdots \lhd t_n.$$

These $n$-saplings capture the growth of an attachment tree up to and including time $n$. Two $n$-saplings $\{t_i\}_{i=1}^n$ and $\{t'_i\}_{i=1}^n$ are considered the same and called isomorphic if $t_i \cong t'_i$ for all $i \in [n]$. We denote the set of (isomorphism classes of) $n$-saplings with weight function $w$ by $\mathcal{T}_{w,n}$ for each $n \in \mathbb{N}$.

Define a family of maps $\phi_n : \mathcal{T}_{w,n} \to \mathcal{R}_n$ by

$$\phi_n\left(\{t_i\}_{i=1}^n\right) := t_n$$

i.e. the map $\phi_n$ simply "forgets" the random recursive structure.

**Definition 1.2.8.** Suppose $t \in \mathcal{R}_n$. Define the *tree multiplicity*,

$$K(t) := \left|\left\{\{t_i\}_{i=1}^n \in \mathcal{T}_{w,n} : \phi_{|t|}^{-1}\left(\{t_i\}_{i=1}^n\right) = t\right\}\right|. \tag{1.2.4}$$

The tree multiplicity, also known as The "Connes-Moscovici weight" [BK99], is defined in [Bro00] and [Sta72, Sect. 22] and is a standard combinatorial function.

## 1.3 Random recursive trees and random Fibonacci sequences

In [MA06], after careful numerical calculation, MacArthur and Anderson remark that it is possible that $\zeta(t_n)^{\frac{1}{n}} \xrightarrow{\text{a.s.}} \mathcal{V}$, given $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$.

Suppose $t \in \mathcal{R}$. In Section 2.1.2 we define two subgroups,

$$\mathcal{E}(t), \mathcal{C}(t) \leq \mathrm{Aut}(t)$$

called the *elementary* and the *non-elementary* subgroups respectively such that

$$|\mathrm{Aut}(t)| = |\mathcal{E}(t)||\mathcal{C}(t)|.$$

This begs the question: does the order of either the elementary or the non-elementary subgroup dominate the other in terms of order? MacArthur further predicted that

$$\begin{aligned} |\mathcal{E}(t_n)|^{\frac{1}{n}} &\xrightarrow{\text{a.s.}} \mathcal{V}, \text{ and} \\ |\mathcal{C}(t_n)|^{\frac{1}{n}} &\xrightarrow{\text{a.s.}} 1 \end{aligned}$$

given $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$.

We remark that there exists $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$ such that

$$\lim_{n\to\infty} \zeta(t_n)^{\frac{1}{n}} \neq \mathcal{V}.$$

For example, consider the random recursive tree $\{t_n\}_{n\in\mathbb{N}}$ where each $t_n$ consists only of a root vertex adjacent to $n-1$ vertices. In this case $\mathrm{Aut}(t_n) \cong S_{n-1}$ since $\zeta(t_n) = (n-1)!$ for all $n \in \mathbb{N}$.

**Remark 1.3.1.** As an aside, the author is not aware of any existing calculation of the expected order (or the almost sure order) of the automorphism group of either the Watts and Strogatz model or the Barabási-Albert model for random graphs.

## 1.4 Results

In Chapter 2 we investigate a geometric decomposition of the automorphism group of a rooted tree into its constituent subgroups. This geometric decomposition turns out to be extraordinarily useful.

In Chapter 3 we use the geometric decomposition to show that we can calculate the limiting order of the automorphism group in terms of the relative abundance of particular subgraphs. We prove that a random recursive tree can be modelled by a family of continuous-time Markov chains called *Crump-Mode-Jagers (C-M-J) processes*. We use the powerful machinery of Branching Processes and the geometric decomposition to prove our first major result, namely Theorem 1.

**Theorem 1.** Let $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$ and $L \in \mathbb{R}$ be such that, $|\mathcal{C}(t_n)|^{\frac{1}{n}} \xrightarrow{\text{a.s.}} L$. Then $L > 1$.

In Chapter 4 we describe in more detail the C-M-J processes that correspond to attachment trees. We apply the machinery of C-M-J processes to prove Theorem 2. Define

$$\mathcal{W} := \exp\left(\sum_{t\in\mathcal{R}} \frac{K(t)\log(\Lambda(t))}{(|t|+1)!}\right).$$

**Theorem 2.** Let $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$. Then,

$$\zeta(t_n)^{\frac{1}{n}} \xrightarrow{\text{a.s.}} \mathcal{W}. \tag{1.4.1}$$

In Chapter 5 we exploit a bijection between the symmetric group on $n$ elements and $(n+1)$-saplings. We give results relating to random permutations and use the bijection to derive results about random recursive trees. We use the *cycle indicator polynomial* to prove Theorem 3.

**Theorem 3.** $\mathcal{W} \neq \mathcal{V}$.

In particular, the "intriguing relationship" remarked on by MacArthur and Anderson in [MA06] does not exist.

In Chapter 6 we use a generating function argument to prove Theorem 4. We show that the expected value of the automorphism group of a random recursive tree grows exponentially.

**Theorem 4.** The expected order, $b_k$, of the automorphism group of a tree $t \in \mathcal{T}_{\widetilde{w},k}$ is

$$b_k = \frac{1}{\sqrt{3}} \left( \left( \frac{1+\sqrt{3}}{2} \right)^k - \left( \frac{1-\sqrt{3}}{2} \right)^k \right). \tag{1.4.2}$$

We remark that the righthand side of Equation 1.4.2 is redolent of the classical Fibonacci sequence.

# Chapter 2

# Background

In Chapter 1 we introduced the elementary and the non-elementary subgroups of the automorphism group. In this chapter we will define these subgroups and explain their significance. In addition we will describe two Markov processes, first defined in [RTV07], that we will use to calculate the limiting behaviour of attachment trees in Chapters 3-5.

In [MSGA08] MacArthur, Sánchez-García and Anderson define a *geometric decomposition* of the automorphism group of any graph. In Section 2.1 we adapt this definition to describe a *rooted geometric decomposition* for the automorphism group of rooted forests. This geometric decomposition turns out to be extraordinarily useful; in Chapter 3 we use it to show that we can calculate the limiting order of the automorphism group of a rooted tree in terms of the relative abundance of particular subgraphs. The rooted geometric decomposition also enables us to define the elementary and non-elementary subgroups.

The two Markov processes that Rudas, Tóth and Valkó define in [RTV07] are formulated in terms of *rooted ordered trees* rather than attachment trees. Growth of rooted ordered trees is captured by the notion of *historical orderings*. In Section 2.2 we define a rooted ordered tree and historical orderings of rooted ordered trees. In Section 2.3 we introduce a discrete Markov process and a continuous-time Markov process that model the growth of rooted ordered trees. We finish this chapter with the statement of Theorem 2.3.3 (Theorem 1 in [RT08]) which gives the limiting distribution of outdegree and induced rooted ordered trees.

There is no new material in this Chapter. The decomposition of $\zeta(t)$ given in Remark 2.1.1 is given in [SB13], the definition of the rooted geometric decomposition is an elementary adaptation of the geometric decomposition given in [MSGA08]. The elementary and non-elementary subgroups were first suggested by MacArthur. The background, definitions, theorems and examples in Section 2.2 and 2.3 can be found in [RTV07] and [RT08].

In Chapter 3 we will prove that the set of historical orderings of rooted ordered trees on $n$ vertices is in bijection with $\mathcal{T}_{w,n}$. Furthermore, we will show that Theorem 2.3.3 can be applied to attachment trees in a straightforward way.

## 2.1 The automorphism group of a rooted forest

In Section 1.3 we remarked that there is a natural way to split the automorphism group into two subgroups: the *elementary subgroup* $\mathcal{E}(t)$ and the *non-elementary subgroup* $\mathcal{C}(t)$. In this section we provide the necessary background for and define $\mathcal{E}(t)$ and $\mathcal{C}(t)$. We begin this Section by setting up notation. In Remark 2.1.1 we see that $\zeta(t)$ can be expressed in terms of the root permutation function $\Lambda(t)$. In Section 2.1.1 we adapt MacArthur, Sánchez-García and Anderson's [MSGA08] geometric decomposition for rooted forests and in Section 2.1.2 we use this new decomposition define the elementary and complex subgroups.

Suppose $f = \bigsqcup_{i\in I} t_i$ is a rooted forest and $V' \subseteq V(f)$. Suppose $f' = \bigsqcup_{j\in J} t'_j$ is the induced graph on $V'$. Each connected component $t'_j$ is an induced subtree of some $t_i$. Let $r(t'_j) \in V(t'_j)$ be the vertex that lies on the unique shortest path from $v'$ to $r(t_i)$ for all $v' \in V(t'_j)$. The *induced rooted forest* $f'$ of $f$ on $V'$ is the induced graph $\bigsqcup_{j\in J} t'_j$ on $V'$ where each $t'_j$ is rooted at $r(t'_j)$.

Suppose $t = (r, V, E)$ is a rooted tree and $v \in V$. Define $t_{\downarrow v}$ to be the induced rooted tree on the set of vertices $w \in V$ such that $v$ lies on the (unique) shortest path from $w$ to $r$ and $d(r, w) \geq d(r, v)$. We say that $t_{\downarrow v}$ is the *progeny* of $t$ at $v$.

Given rooted tree $t = (r, V, E)$ we write $B^-(t)$ for the induced rooted forest on $V - \{r\}$. We write $B^-(\bullet) = 1$ and $B^+(1) = \bullet$. For example, let $f$ and $t$ be as in the statement of Definition 1.2.1, then $f = B^-(t)$. Given a rooted tree $t \in \mathcal{R}$,

$$B^+(B^-(t)) = t = B^-(B^+(t)).$$

In [Jor69] Jordan observed that the automorphism group of a graph is determined by the automorphism group of its connected components and used this to characterise the class of groups that arise from automorphism groups of rooted forests.

Suppose $f$ is a rooted forest that consists of two non-isomorphic rooted trees $t_1$ and $t_2$. Every automorphism of $f$ is obtained via an automorphism of $t_1$ along with an automorphism of $t_2$. Now suppose $f_1$ is a rooted forest that consists of $m_1$ copies of $t_1$. Automorphism of $f$ are obtained by automorphisms of individual copies of $t_1$ in addition to a permutation of those copies. Putting this together, let rooted forest $f$ be as in Definition 1.2.1. Then,

$$\mathrm{Aut}(f) \cong (\mathrm{Aut}(t_1) \wr S_{\alpha_1}) \times \cdots \times (\mathrm{Aut}(t_k) \wr S_{\alpha_k}). \tag{2.1.1}$$

Equation 2.1.1 adapted for unrooted graphs can be found in [Har69].

Suppose $t = B^+(f)$. It is clear that $\mathrm{Aut}(t) \cong \mathrm{Aut}(f)$.

**Remark 2.1.1.** The order of the automorphism group of a rooted tree can be given in terms of the root permutation function as follows.

Suppose $t = B^+(t_1^{\alpha_1}, \ldots, t_m^{\alpha_m})$. Then, $\mathrm{Aut}(t) \cong (\mathrm{Aut}(t_1) \wr S_{\alpha_1}) \times \cdots \times (\mathrm{Aut}(t_m) \wr S_{\alpha_m})$. Hence,

$$\zeta(t) = \prod_{i=1}^{m} \zeta(t_i)^{\alpha_i}(\alpha_i!). \tag{2.1.2}$$

Furthermore, $\zeta(t) = \prod_{v \in V(t)} \Lambda(t_{\downarrow v})$ by induction.

### 2.1.1 The rooted geometric decomposition

In [MSGA08] MacArthur, Sánchez-García and Anderson prove that there exists a *geometric* decomposition of the automorphism group of a graph into a direct product of *geometric factors*. In this section we adapt their geometric decomposition for rooted forests.

Suppose that $S_n$ acts on a set $X$ and that $\sigma \in S_n$. We define the *support* of $\sigma$ to be the collection of elements of $X$ that are not fixed by $\sigma$. We write

$$\mathrm{supp}(\sigma) = \{x \in X : \sigma(x) \neq x\}$$

Two permutations are said to be *disjoint* if their supports are non-intersecting. Two sets of permutations $P$ and $Q$ are *support-disjoint* if every pair of permutations $\sigma \in P$ and $\tau \in Q$ have disjoint supports.

Let $f$ be a rooted forest and $\mathcal{S}$ be a minimal set of generators for $\mathrm{Aut}(f)$. Partition $\mathcal{S}$ into $m$ support-disjoint subsets $\mathcal{S}_1 \cup \mathcal{S}_2 \cup \cdots \cup \mathcal{S}_m$ such that each $\mathcal{S}_i$ cannot be partitioned into smaller support-disjoint subsets. Let $H_i$ be the subgroup of $\mathrm{Aut}(f)$ generated by $\mathcal{S}_i$. Following MacArthur, Sánchez-García and Anderson we say that each $H_i$ is a *rooted geometric factor* of $\mathrm{Aut}(f)$.

Since $\mathcal{S}$ generates $\mathrm{Aut}(f)$ and any pair $H_i$ and $H_j$ (where $i \neq j$) commute there is a direct product decomposition,

$$\mathrm{Aut}(f) \cong H_1 \times H_2 \times \cdots \times H_m, \tag{2.1.3}$$

that we call the *rooted geometric decomposition.* In [MSGA08, Proposition 2.1] MacArthur, Sánchez-García and Anderson prove that the geometric decomposition is well-defined, i.e. it does not depend on the choice of generating set and the geometric factors cannot themselves be written as $K \times L$ with $K$ and $L$ support-disjoint subgroups. We remark that [MSGA08, Proposition 2.1] applies to any finite permutation group, not just automorphism groups of graphs. In particular the rooted geometric decomposition is well-defined.

Following MacArthur, Sánchez-García and Anderson we define a *rooted symmetric motif* to be the induced rooted forest on the support of a rooted geometric factor $H$.

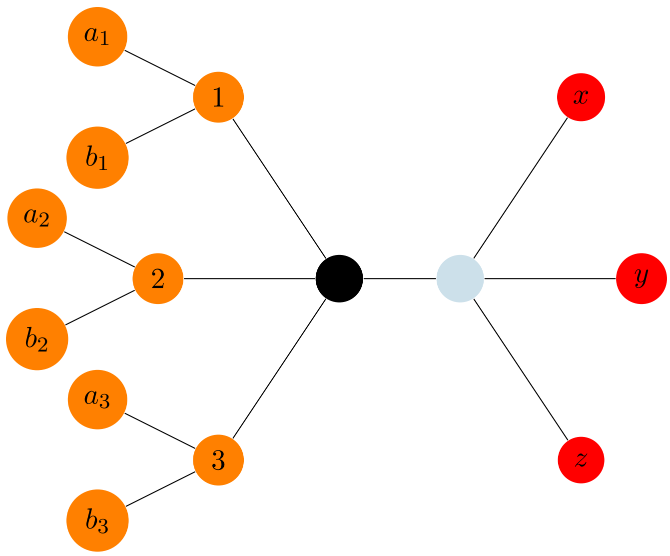


Figure 2.1: A rooted tree $t \in \mathcal{R}$. We have indicated the root by shading $r(t)$ black.

**Example 2.1.2.** Consider the rooted tree $t$ in Figure 2.1. Let,

$$X_1 = \{(x, y), (y, z)\}, \quad \text{and} \quad X_2 = \{(1, 2)(a_1, a_2)(b_1, b_2), (2, 3)(a_2, a_3)(b_2, b_3), (a_1, b_1)\}$$

Clearly $X_1 \cup X_2$ generate $\mathrm{Aut}(t)$. Since

$$\mathrm{supp}(X_1) = \{x, y, z\} \text{ and } \mathrm{supp}(X_2) = \{1, 2, 3, a_1, a_2, a_3, b_1, b_2, b_3\},$$

$X_1$ and $X_2$ are support-disjoint. Neither $X_1$ or $X_2$ can be partitioned into smaller support-disjoint subsets. Since $X_1$ generates the symmetric group $S_3$ and $X_2$ generates $S_2 \wr S_3$,

$$\mathrm{Aut}(t) \cong S_3 \times (S_2 \wr S_3)$$

is the rooted geometric decomposition of $\mathrm{Aut}(t)$.

The induced rooted forest on the red vertices $\{x, y, z\} \in V(t)$ is the rooted symmetric motif corresponding to the rooted geometric factor $S_3 \leq \mathrm{Aut}(t)$. The induced rooted forest on the orange vertices $\{1, 2, 3, a_1, a_2, a_3, b_1, b_2, b_3\} \in V(t)$ is the rooted symmetric motif corresponding to the rooted geometric factor $S_2 \wr S_3 \leq \mathrm{Aut}(t)$.

### 2.1.2 The elementary and non-elementary subgroups

A $(n, k)$-*star* is a rooted tree consisting of a root vertex that is adjacent to exactly $k$ paths of length $n$ and no other vertices for some $n, k \in \mathbb{N}$. For succinctness we will refer to $(1, k)$-stars simply as $k$-*stars*.

Suppose $t$ is a rooted tree and, via the rooted geometric decomposition, $\mathrm{Aut}(t) = \prod_{i=1}^{m} H_i$. Let $G_i$ be the rooted symmetric motif corresponding to $H_i$ for each $i \in [m]$. We partition $[m]$ into two sets as follows,

$$A := \{i \in [m] : H_i \cong B^-(s)\} \quad B := [m] - A$$

where $s$ is a $(n, k)$-star. We define the *elementary subgroup* of $\mathrm{Aut}(t)$ to be,

$$\mathcal{E}(t) = \prod_{i \in A} H_i$$

and the *non-elementary subgroup*,

$$\mathcal{C}(t) = \prod_{i \in B} H_i.$$

Loosely speaking the elementary subgroup captures the contribution to the automorphism group coming from $(n, k)$-stars and the non-elementary subgroup captures the contribution to the automorphism group coming from more complicated induced rooted forests.

## 2.2 Rooted ordered trees

In this section we will define and investigate a family of trees called *rooted ordered trees* which will allow us to define two Markov processes in Section 2.3. There is no new material in this section; it closely follows Section 1.1 and 1.3 of [RT08]. In Section 2.2.1 we investigate the growth of rooted ordered trees via *historical orderings* which we can think of like $n$-saplings. We go on, in Section 2.2.2, and use the weight function $w : \mathbb{N} \to \mathbb{R}^>$ to define the *historical sequence of weights* for a rooted ordered tree.

**Definition 2.2.1.** [RTV07, Ner81].

$$\mathcal{N} := \bigsqcup_{n=0}^{\infty} \mathbb{N}^n$$

where $\mathbb{N}^0 := \{\emptyset\}$ .

A typical element $x \in \mathcal{N}$ has the form $x = (x_1, x_2, \ldots, x_n)$ where $x_1, \ldots, x_n \in \mathbb{N}$. Suppose $x = (x_1, x_2, \ldots x_n) \in \mathcal{N}$ and $y = (y_1, y_2, \ldots y_k) \in \mathcal{N}$. We write $xy$ for the concatenation

$$xy = (x_1, x_2, \dots x_n, y_1, y_2, \dots y_k) \in \mathcal{N}. \tag{2.2.1}$$

Suppose $t$ is a rooted tree. We arbitrarily assign a total order $<_v$ to $\mathrm{Ch}_v$ for every $v \in V(t)$. Suppose $\mathrm{Ch}_v = \{w_1, w_2, \dots, w_k, \dots\}$ and

$$w_1 <_v w_2 <_v \cdots <_v w_k$$

then we say that $w_i$ is the $i^{th}$ child of $v$.

**Definition 2.2.2.** (rooted ordered tree) A *rooted ordered tree* $G = (r, V, E, <, l)$ is a rooted tree together with a family of total orderings, $< := \{<_v\}_{v \in V}$, on each set of children $\mathrm{Ch}_v$ and a labelling $l : V \to \mathcal{N}$. This labelling is defined by $l(r) := \emptyset$ and then recursively by,

$$l(w_i) := l(v)i$$

(with the concatenation given by Equation 2.2.1) whenever $w_i$ is the $i^{th}$ child of $v$.

**Remark 2.2.3.** The vertices of a rooted ordered tree are thus labelled by elements of $\mathcal{N}$ as follows: $\emptyset$ denotes the root of the rooted ordered tree. The first child of the root vertex is labelled 1, the second child 2, the third child 3 and so on. In general the vertex labelled $(x_1, x_2, \dots x_n)$ is the $x_n^{th}$ child of the $x_{n-1}^{th}$ child of the ... of the $x_1^{th}$ child of the root vertex.

We denote the set of all rooted ordered trees on $n$ vertices $\mathcal{G}_n$. We have drawn $\mathcal{G}_4$ in Figure 2.2.

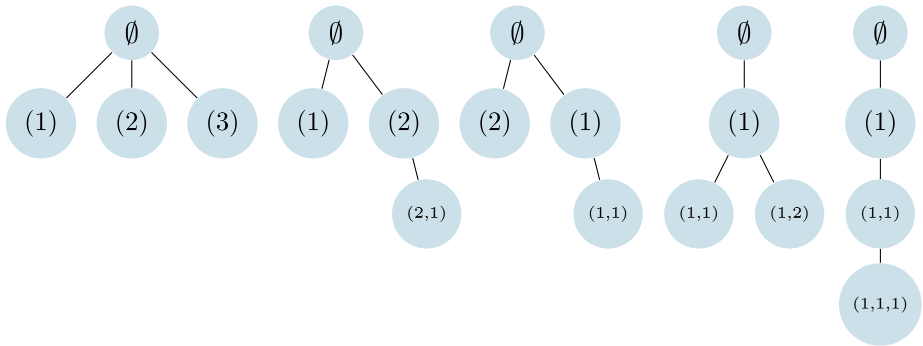


Figure 2.2: We have listed all elements of $\mathcal{G}_4$. We remark that $|\mathcal{G}_4| = 5$.

Define

$$\mathcal{G} = \bigcup_{n \in \mathbb{N}} \mathcal{G}_n.$$

Since $\mathcal{G}_n$ contains a finite number of rooted labelled trees for each $n \in \mathbb{N}$ the set $\mathcal{G}$ is countable.

**Remark 2.2.4.** Following [RTV07] we identify a rooted ordered tree with the set of labels of its vertices since this contains all the necessary information about its edges.

We shall therefore abuse notation and write $x \in G$ for a vertex $x$ of a rooted ordered tree $G$.

### 2.2.1 Historical orderings

Let $G \in \mathcal{G}$. An ordering $s = (s_1, \ldots, s_{|G|})$ of the vertices of $G$ is called *historical* if

$$\{s_1, \ldots, s_i\} \in \mathcal{G}$$

for each $1 \leq i \leq |G|$. The *set of all historical orderings of* $G \in \mathcal{G}$ is denoted $S(G)$ [RTV07].

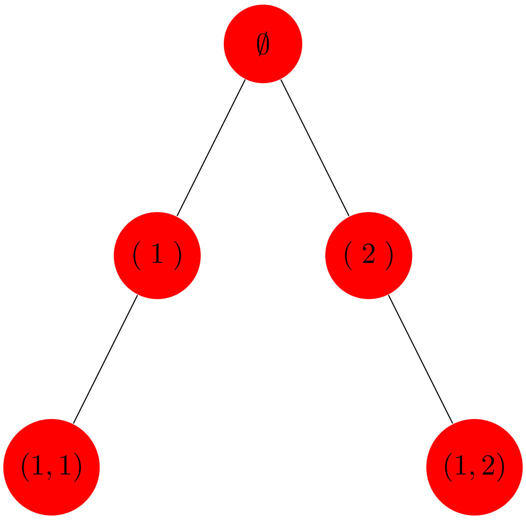


Figure 2.3: A rooted ordered tree $G \in \mathcal{G}$ such that $G$ has vertices $\{\emptyset, (1), (2), (1,1), (1,2)\}$. There are three distinct historical orderings for $G$ which are $(\emptyset, (1), (1,1), (2), (2,1))$, $(\emptyset, (1), (2), (1,1), (2,1))$, and $(\emptyset, (1), (2), (2,1), (1,1))$.

We denote the *set of historical orderings of rooted ordered trees* on $n$ vertices, $\mathcal{SG}_n$. For a fixed rooted ordered tree $G \in \mathcal{G}$ and $s \in S(G)$, the sequence of rooted ordered trees,

$$G(s,i) := \{s_1, \ldots, s_i\}$$

for $i = 1, 2, \ldots, |G|$ is called the *evolution* of $G$ in the historical ordering $s$. We should think of the evolution of $G$ analogously to the $n$-saplings defined in Section 1.2.2.

**Remark 2.2.5.** [Section 2.1][RTV07] A sequence $s = \{s_1, \ldots, s_{|G|}\}$ with each $s_i \in \mathcal{N}$ is a historical ordering of a rooted ordered tree $G$ if and only if

(H1) $s_1 = \emptyset$ and $s_2 = (1)$,

(H2) if $s_j = (x_1, x_2, \ldots, x_m)$ then there exists $k < j$ such that $s_k = (x_1, x_2, \ldots, x_{m-1})$, and

(H3) if $s_j = (x_1, x_2, \ldots, x_m)$ and $x_m \geq 2$ then there exists $k < j$ such that $s_k = (x_1, x_2, \ldots, x_m - 1)$.

### 2.2.2 Weight function

Suppose $w : \mathcal{N} \to \mathbb{R}^{>}$ is a weight function. The *total weight* of a rooted ordered tree is

$$W_w(G) := \sum_{x \in G} w(\mathrm{out}(x, G)).$$

Fix a historical ordering $s = (s_1, \ldots, s_{|G|})$ for some $G \in \mathcal{G}$. Define the *historical sequence of total weights* to be

$$W_w(G, s, i) := W_w(G(s, i))$$

for each $1 \leq i \leq |G|$ [RTV07].

**Example 2.2.6.**

$$\begin{aligned} W_{\widetilde{w}}(G) &= \sum_{x \in G} \widetilde{w}(\mathrm{out}(x, G)) \\ &= \sum_{x \in G} 1 \\ &= |G|. \end{aligned}$$

Thus for a fixed historical ordering $s = (s_1, \ldots, s_{|G|})$, the historical sequence of total weights is given by $W_{\widetilde{w}}(G, s, i) = i$ for each $1 \leq i \leq |G|$.

Given the historical sequence of total weights it makes sense to define a historical sequence of weights for each vertex in a rooted ordered tree. Suppose $s = (s_1, s_2, \ldots, s_{|G|})$ is a historical ordering of the vertices of a rooted ordered tree $G$ and the parent of vertex $s_i$ is $p(s_i)$ for each $i = 2, 3, \ldots, |G|$. The weights of each vertex $s_i$ are given by

$$w(G, s, i) := w(\mathrm{out}(p(s_i), G(s, i-1))$$

for each $2 \leq i \leq |G|$ [RTV07]. For example, $\widetilde{w}(G, s, i) = 1$ for all $2 \leq i \leq |G|$.

## 2.3 Two models of randomly growing trees

Rudas, Tóth and Valkó define a continuous and a discrete model of randomly growing trees in [RTV07] that we will describe in this section. The parameter of these models is the weight function. To ensure the continuous model is well-defined Rudas, Tóth and Valkó place a restriction on the weight function that we give in Section 2.3.1.

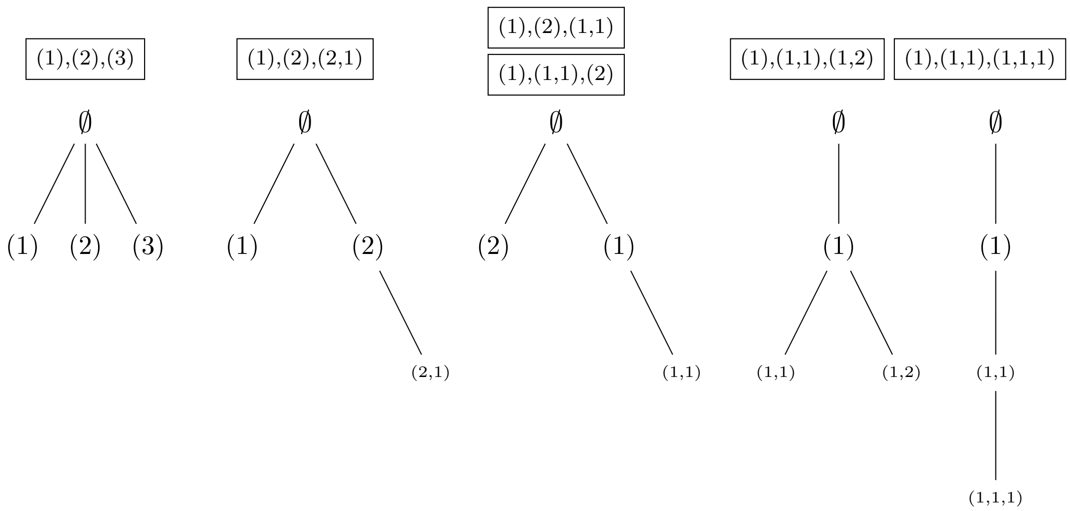


Figure 2.4: We have listed all elements of $\mathcal{SG}_4$ above the rooted ordered tree for which they are a historical ordering. For example, there is precisely one historical labelling of the rooted ordered tree on the far left namely $\{(1), (2), (3)\}$. We remark that $|\mathcal{SG}_4| = 6$.

We do not give the definitions of elementary concepts from branching processes and Markov chains such as *point processes*, *jump chains*, *pure birth processes*, *holding times*, *jump times* and *jump processes*. We recommend [Nor98] for a reference.

The most important result of this Section is Theorem 2.3.3 part (ii) (Theorem 1 in [RT08]) which gives the limiting distribution of induced subtrees of a randomly growing rooted ordered tree. In Chapter 3 we will adapt Theorem 2.3.3 for random recursive trees (see Proposition 3.2.1).

Rudas, Tóth and Valkó [RTV07], define a discrete time Markov chain $\mathcal{Y}_w^d$ on the countable space $\mathcal{G}$ with initial state $\mathcal{Y}_w^d(0) = \{\emptyset\}$. If, for $n \geq 0$ we have $\mathcal{Y}_w^d(n) = G$ then for a vertex $x \in V(G)$ let $k = \text{out}(x, G) + 1$. The transition probabilities are:

$$\mathbb{P}\left(\mathcal{Y}_w^d(n+1) = G \cup \{xk\}\right) = \frac{w(\text{out}(x, G))}{\sum_{y \in G} w\left(\text{out}(y, G)\right)}.$$

So at each step a new vertex is attached via an edge to exactly one existing vertex. Suppose $p : n \mapsto \mathcal{Y}_w^d(n)(\omega)$ is a path of process $\mathcal{Y}_w^d(n)$ such that if, for $n \geq 0$, we have $\mathcal{Y}_w^d(n)(\omega) = G_n$ then

$$G_{n+1} := G_n \cup \{x_n k_n\}$$

where $x_n \in V(G_n)$ and $k_n = \text{out}(x_n, G_n) + 1$. Rudas, Tóth and Valkó [RTV07] go on to define a continuous Markov process $\mathcal{Y}_w(\tau)_{\tau \in \mathbb{R}}$ with state space $\mathcal{G}$ (identified with certain subsets of $\mathcal{N}$ as in Remark 2.2.4). The initial distribution $\mathcal{Y}_w(0) = \emptyset$. Again suppose that $G \in \mathcal{G}$, fix some vertex $x \in G$ and let $k = \text{out}(x, G) + 1$. If $\mathcal{Y}_w(\tau) = G$ then the process may jump to $G \cup \{xk\}$ with rate $w(\text{out}(x, G))$. This means that each existing vertex $x \in \mathcal{Y}_w(\tau)$ "gives birth" to a child vertex with rate $w(\mathrm{out}(x, \mathcal{Y}_w(\tau)))$ independently of other vertices.

### 2.3.1 A restriction on the weight function

In order that $\mathcal{Y}_w(\tau)$ is well-defined Rudas, Tóth and Valkó [RTV07], put a restriction on the weight function $w$ that we will define in this section.

Let $X_w(\tau)$ be the Markovian pure birth process defined by $X_w(0) = 0$ and birth rates

$$\mathbb{P}\left(X_w(\tau + d\tau) = k + 1 | X_w(\tau) = k\right) = w(k)d\tau + o(d\tau).$$

The holding times $H_1$,$H_2, \dots$ are independent exponentially distributed random variables of parameters $w(0), w(1), w(2), \dots$ and with jump chain $Y_n = n$ for every $n \in \mathbb{N}$ [Nor98]. Let $J_{w,1}, J_{w,2}, \dots$ be the jump times associated with $X_w(\tau)$.

We associate a point process $\xi_w := (\xi_{w,1}, \xi_{w,2}, \dots)$ with $X_w(\tau)$ in the usual way. We define $\xi_{w,i} = J_{w,i}$ for each $i \in \mathbb{N}$. In this case the $\xi_w$-measure, $\xi_w(\tau) := \xi_w([0, \tau])$ is given by

$$\xi_w(\tau) = \mathrm{out}(\emptyset, \mathcal{Y}_w(\tau)),$$

which is the random number of individuals born up to time $\tau$.

Suppose that $\rho_w$ is the density of the point process $\xi_w$ . Then,

$$\widehat{\rho}_w(\lambda) = \int_0^\infty e^{-\lambda\tau} \rho_w(\tau) d\tau$$

is the *Laplace transform* of $\rho_w$.

Let $\underline{\lambda}_w := \inf\{\lambda > 0 : \widehat{\rho}_w(\lambda) < \infty\}$. Throughout the remainder of this thesis we impose the following condition on the weight function:

$$\lim_{\lambda \to \underline{\lambda}_w} \widehat{\rho}_w(\lambda) > 1. \tag{2.3.1}$$

Rudas, Tóth and Valkó prove that the equation $\widehat{\rho}_w(\lambda) = 1$ has a unique root for all $\lambda \in \mathbb{R}$. We denote this unique root $\lambda_w^*$.

**Example 2.3.1.** Suppose that $X_{\widetilde{w}}(t)$ is the Markovian pure birth process with $X_{\widetilde{w}}(0) = 0$ and birth rates $(j_i : i \geq 0)$. Rudas and Tóth [RT08, Section 1.2.2] remark that

$$\widehat{\rho}(\lambda) = \sum_{n=1}^{\infty} \prod_{i=0}^{n-1} \frac{j_i}{\lambda + j_i}.$$

Suppose that $X_{\widetilde{w}}(t)$ is the Markovian pure birth process with $X_{\widetilde{w}}(0) = 0$ and birth rates $j_i = 1$. Then,

$$\begin{aligned}
\widehat{\rho}_{\widetilde{w}}(\lambda) &= \sum_{n=1}^{\infty} \prod_{i=0}^{n-1} \frac{j_i}{\lambda + j_i} \\
&= \sum_{n=1}^{\infty} \prod_{i=0}^{n-1} \frac{1}{\lambda + 1} \\
&= \sum_{n=1}^{\infty} \left( \frac{1}{\lambda + 1} \right)^n \\
&= \frac{1}{\lambda}
\end{aligned}$$

by the theory of geometric series. Therefore $\lambda^*_{\widetilde{w}} = 1$. We remark that $\underline{\lambda}_{\widetilde{w}} = 0$, hence $\widetilde{w}$ satisfies Condition 2.3.1.

Recall that given a rooted tree $t \in \mathcal{R}$ we denote the induced subtree rooted at $v \in V(t)$ by $t_{\downarrow v}$. Analogously we write $G_{\downarrow x}$ for the *induced rooted ordered subtree* of $G$ rooted at $x \in G$. Furthermore, we write $\mathcal{Y}_{\downarrow x}(\tau)$ for the induced subtree of $\mathcal{Y}(\tau)$ rooted at $x$ i.e the set of descendants of $x$ (including $x$) born before and including time $\tau$.

**Definition 2.3.2.** [RTV07]. Given a weight function $w$ satisfying condition 2.3.1, $k \in \mathbb{N}$ and $G \in \mathcal{G}$, define

$$p_w(k) := \frac{\lambda^*_w}{\lambda^*_w + w(k)} \prod_{i=0}^{k-1} \frac{w(i)}{\lambda^*_w + w(i)}, \tag{2.3.2}$$

$$\pi_w(G) := \sum_{s \in S(G)} \frac{\lambda^*_w}{\lambda^*_w + W_w(G)} \prod_{i=0}^{|G|-2} \frac{w(G, s, i+2)}{\lambda^*_w + W_w(G, s, i+1)}. \tag{2.3.3}$$

We are now ready to state a theorem from [RT08].

**Theorem 2.3.3.** [RT08, Theorem 1] Let $w$ be a weight function that satisfies Condition 2.3.1. The following limits hold almost surely:

(i) For any fixed $k \in \mathbb{N}$,

$$\lim_{\tau \to \infty} \frac{|\{x \in \mathcal{Y}_w(\tau) : \mathrm{out}(x, \mathcal{Y}_w(\tau)) = k\}|}{|\mathcal{Y}_w(\tau)|} = p_w(k).$$

(ii) Fix $G \in \mathcal{G}$, then

$$\lim_{\tau \to \infty} \frac{|\{x \in \mathcal{Y}_w(\tau) : \mathcal{Y}_w(\tau)_{\downarrow x} = G\}|}{|\mathcal{Y}_w(\tau)|} = \pi_w(G).$$

**Example 2.3.4.** This example is a subcase of those dealt with by Rudas and Tóth in [RTV07] who deal with a linear weight function in [RT08, Section 2.2]. Suppose $k \in \mathbb{N}$ and $G \in \mathcal{G}$. Then,

(i) By Example 2.3.1 $\lambda^*_{\widetilde{w}} = 1$. Hence,

$$\begin{aligned} p_{\widetilde{w}}(k) &= \frac{1}{1+1} \prod_{i=0}^{k-1} \frac{1}{1+1} \\ &= \frac{1}{2} \left(\frac{1}{2}\right)^k \\ &= \frac{1}{2^{k+1}}. \end{aligned}$$

(ii) By Example 2.2.6 $W_{\widetilde{w}}(G) = |G|$ and $W_{\widetilde{w}}(G, s, i) = i$ for all $G \in \mathcal{G}$, $s \in S(G)$ and $i \in [|G|]$. Therefore,

$$\begin{aligned} \pi_{\widetilde{w}}(G) &= \sum_{s \in S(G)} \frac{1}{1+|G|} \prod_{i=0}^{|G|-2} \frac{1}{1+(i+1)} \\ &= \sum_{s \in S(G)} \frac{1}{(|G|+1)!} \end{aligned}$$

for any $G \in \mathcal{G}$.

# Chapter 3

# Two models of random attachment trees

In Chapter 2 we investigated two Markov processes first defined in [RTV07] that model the growth of rooted ordered trees. This growth is captured formally in the notion of *historical sequences*. Rudas, Tóth and Valkó [RTV07] used these models to investigate the limiting properties of historical orderings of rooted ordered trees.

In this chapter we will construct a bijection $B : \mathcal{SG}_n \to \mathcal{T}_{w,n+1}$ and show that we may adapt Theorem 2.3.3 for random attachment trees. We use this to prove our first major result which is Theorem 1. We remark that although bijection $B$ is obvious it is our own work and not from the literature.

In order to construct $B : \mathcal{SG}_n \to \mathcal{T}_{w,n+1}$, in Section 3.1 we show that every $n$-sapling $\{t_i\}_{i=1}^n$ can be associated with a vertex sequence, $\mathbf{v}\left(\{t_i\}_{i=1}^n\right)$, that encodes the vertices chosen for attachment in the construction of $\{t_i\}_{i=1}^n$. We use this to construct $B$. Then, in Section 3.2 we prove that $\mathcal{Y}_w^d$ is a model for random attachment trees with weight function $w$.

The rooted geometric decomposition of the automorphism group of a tree described in Section 2.1 allows us to calculate the automorphism group of a tree via the abundance of particular rooted subtrees called *rooted symmetric motifs* introduced in Section 2.1.1. In Section 3.2 we put this together with Theorem 2.3.3 adapted for random recursive trees to prove our first major result; namely Theorem 1.

The vertex sequence associated with a $n$-sapling is just a subtle variant on a Prüfer sequence (see [AZHE10] or [Bol13] for a definition). The existence of bijection $B$ is obvious and its construction elementary. The proof of Theorem 1 is also straightforward.

## 3.1 Correspondence between saplings and historical sequences

In this section we will prove that there exists a correspondence between $n$-saplings and historical orderings of rooted ordered trees. This enables us to adapt Theorem 2.3.3 for random recursive trees.

Suppose $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_w$. We recursively define a sequence $\{l_n\}_{n\in\mathbb{N}}$ of labellings $l_n : V(t_n) \to [n]$ by $l_1(\bullet) = 1$ and subsequently

$$l_n(v) = \begin{cases} l_{n-1}(v) & \text{if } v \in V(t_{n-1}) \\ n & \text{otherwise} \end{cases}$$

This gives an immediate correspondence between $n$-saplings $\{t_i\}_{i=1}^n$ and the rooted tree $t_n$ together with the labelling $l_n$.

For example suppose $t_1, t_2, t_3$ are the rooted trees shown in Figure 3.1. We represent the random attachment tree $t_1 \lhd t_2 \lhd t_3$ by the rooted tree and labelling shown in Figure 3.2.

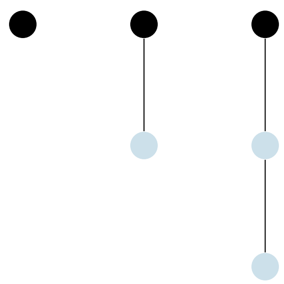

Figure 3.1: Rooted trees $t_1 \in \mathcal{R}_1$, $t_2 \in \mathcal{R}_2$ and $t_3 \in \mathcal{R}_3$.


Figure 3.2: The labelled rooted tree that corresponds to $t_1 \lhd t_2 \lhd t_3 \in \mathcal{T}_{w,3}$.

Every $n$-sapling can be associated with a sequence

$$\mathbf{v}\left(\{t_i\}_{i=1}^n\right) := (v_2, v_3, \ldots, v_n).$$

where $v_i = l_n(u_{i-1})$, i.e. the label of the vertex chosen for attachment at time $i$. For example the 5-sapling $\{t_i\}_{i=1}^5 \in \mathcal{T}_{w,5}$ in Figure 3.3 has vertex sequence $\mathbf{v}(\{t_i\}_{i=1}^5) = (1,1,2,3)$. Suppose that $\mathbf{v}(\{t_i\}_{i=1}^n)$ is a vertex sequence and fix some $k \in \mathbb{N}$. Define $d_k(\{t_i\}_{i=1}^n)$ to be the number of occurrences of $k$ in $\mathbf{v}(\{t_i\}_{i=1}^n)$. Suppose $v \in V(t_n)$ and $l_n(v) = k$. Then,

$$\operatorname{out}(v, t_n) = d_k(\{t_j\}_{j=1}^n)$$

for all $k \in [n]$.

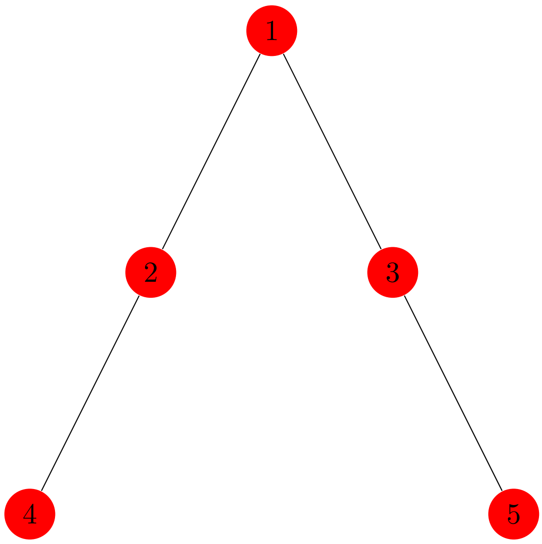


Figure 3.3: The 5-sapling $\{t_i\}_{i=1}^5 \in \mathcal{T}_{w,5}$. The associated vertex sequence is $\mathbf{v}(\{t_i\}_{i=1}^5) = (1,1,2,3)$.

In Figure 2.4 we noted that $|\mathcal{SG}_4| = 6$. Since we also have $|\mathcal{T}_{w,4}| = 6$ this points to a correspondence. In this section we will show that there exists a bijection between $\mathcal{T}_n$ and $\mathcal{SG}_n$ by constructing two functions,

$$g : \mathcal{T}_{w,n} \longrightarrow \mathcal{SG}_n$$
$$h : \mathcal{SG}_n \longrightarrow \mathcal{T}_{w,n}$$

that we will show to be inverse.

Fix a $n$-sapling $\{t_i\}_{i=1}^n \in \mathcal{T}_{w,n}$ and let $\mathbf{v}(\{t_i\}_{i=1}^n) = (v_2, v_3, \ldots, v_n)$. Define a function $\tilde{g}_n : [n] \to \mathcal{N}$ by setting $\tilde{g}_n(1) := \emptyset$, then recursively

$$\tilde{g}_n(i) := \tilde{g}_n(v_i) d_{v_i}(\{t_j\}_{j=1}^i)$$

for $i = 2, 3, 4, \ldots, n$.

Now we define a map $g : \mathcal{T}_{w,n} \to \mathcal{SG}_n$ as follows:

$$g(\{t_i\}_{i=1}^n) = (\tilde{g}_n(1), \tilde{g}_n(2), \ldots, \tilde{g}_n(n)) \tag{3.1.1}$$

for any $\{t_i\}_{i=1}^n \in \mathcal{T}_{w,n}$.

**Lemma 3.1.1.** Suppose $\{t_i\}_{i=1}^n \in \mathcal{T}_{w,n}$. Then $g(\{t_i\}_{i=1}^n)$ is a historical ordering for a rooted ordered tree $G \in \mathcal{G}_n$.

*Proof.* Since $g : \mathcal{T}_{w,n} \to \mathcal{N}$, by Proposition 2.2.5 it is enough to check conditions (H1) - (H3). Let $\mathbf{v}\left(\{t_i\}_{i=1}^n\right) = (v_2, v_2, \ldots, v_n)$. By the construction of map $g$ we have $s_j = \tilde{g}_n(j)$.

Condition $(H1)$ is trivially satisfied. Suppose $s_j = (x_1, x_2, \ldots, x_m)$ and $m > 1$. There exists a vertex $v \in V(t_n)$ labelled $l_n(v) = j$. By the construction of the vertex sequence associated with a $n$-sapling the parent $w \in V(t_n)$ of $v$ is labelled $l_n(w) = v_j$.

By the construction of the historical ordering (see Section 2.2.1) $s_j = \tilde{g}_n(j)d_{v_j}(\{t_i\}_{i=1}^j)$, hence $s_{v_j} = (x_1, x_2, \ldots, x_{m-1})$. Condition (H2) is satisfied since $v_j \in [j-1]$.

Suppose $s_j = (x_1, \ldots, x_m)$ and $x_m > 1$. Since $s_j = \tilde{g}_n(j)d_{v_j}(\{t_i\}_{i=1}^j)$ we have $d_{v_j}(\{t_i\}_{i=1}^j) = x_m > 1$. Since $d_{v_j}(\{t_i\}_{i=1}^j) = \mathrm{out}(v_j, t_j) > 1$, there exists a non-empty set $K$ such that for all $k \in K$ we have $k < j$ and $v_k = v_j$. Let $\tilde{k} = \max\{k \in K\}$. By the construction of historical sequences we have $s_{\tilde{k}} = (x_1, \ldots, x_{m-1}, x_m - 1)$, hence (H3) is satisfied. $\square$

**Definition 3.1.2** (ancestor function)**.** Suppose that $x = (x_1, x_2, \ldots, x_m) \in \mathcal{N}$. We define the *ancestor* function $\mathrm{anc} : \mathcal{N}\backslash\{\emptyset\} \to \mathcal{N}$ as follows:

$$\mathrm{anc}(x) = \begin{cases} \emptyset & \text{if } m = 1 \\ (x_1, x_2, \ldots, x_{m-1}) & \text{otherwise} \end{cases}$$

Suppose that $s_G = (s_1, s_2, \ldots, s_n)$ is a historical ordering of a rooted ordered tree $G \in \mathcal{G}$. Define a further function,

$$\tilde{h}(s_i) := i.$$

for all $i \in [n]$. Define $h(s) := (v_2, v_3, \ldots, v_n)$ where each $v_i = \tilde{h}\left(\mathrm{anc}\left(s_i\right)\right)$ for $2 \le i \le n$.

**Lemma 3.1.3.** Suppose $s \in \mathcal{SG}_n$. Then there exists $\{t_i\}_{i=1}^n \in \mathcal{T}_{w,n}$ such that

$$h(s) = \mathbf{v}\left(\{t_i\}_{i=1}^n\right).$$

*Proof.* Suppose $s = (s_1, \ldots, s_n)$. It is enough to prove that $v_i = \tilde{h}\left(\mathrm{anc}\left(s_i\right)\right) \in [i-1]$ for $i = 2, \ldots, n$. By (H2), if $s_i = (x_1, x_2, \ldots, x_m)$ and $m > 1$ then there exists $k \in [i-1]$ such that $s_k = (x_1, x_2, \ldots, x_{m-1})$. We note that in this case $s_k = \mathrm{anc}\left(s_i\right)$. Hence,

$$\begin{aligned} v_i &= \tilde{h}\left(\mathrm{anc}\left(s_i\right)\right) \\ &= \tilde{h}(s_k) \\ &= k. \end{aligned}$$

□

**Theorem 3.1.4.** $\mathcal{T}_{w,n}$ and $\mathcal{SG}_n$ are in bijection.

*Proof.* We prove the lemma by showing that $g$ and $h$ defined above are inverse. Suppose $\{t_i\}_{i=1}^n \in \mathcal{T}_{w,n}$ and that $\mathbf{v}\left(\{t_i\}_{i=1}^n\right) = (v_2, v_3, v_4, \ldots, v_n)$ is the associated vertex sequence and $g\left(\{t_i\}_{i=1}^n\right) = s$ where $s = (s_1, \ldots, s_n)$.

We begin by remarking that at time $i$ in the construction of the $n$-sapling $v_i$ is the label of the *parent* of vertex $i$. Hence $d_{v_i}\left(\{t_j\}_{j=1}^i\right) \geq 1$ and

$$\operatorname{anc}\left(\tilde{g}_n(v_i)d_{v_i}\left(\{t_j\}_{j=1}^i\right)\right) = \tilde{g}_n(v_i) \tag{3.1.2}$$

for $i = 2, 3, \ldots n$. By Lemma 3.1.1 there exists $s = (s_1, s_2, \ldots, s_n) \in \mathcal{SG}_n$ such that $g(t_n) = s$. By the definition of $g$ we have $s_i = \tilde{g}_n(i)$ for $i = 2, 3, \ldots, n$ (see Equation 3.1.1). Therefore, $\tilde{h}\left(\tilde{g}_n(i)\right) = i$ and in particular

$$\tilde{h}\left(\tilde{g}_n\left(v_i\right)\right) = v_i \tag{3.1.3}$$

for $i = 2, 3, \ldots, n$. Putting Equation 3.1.2 and Equation 3.1.3 together we get

$$\begin{aligned}\tilde{h}\left(\operatorname{anc}\left(\tilde{g}_n(i)\right)\right) &= \tilde{h}\left(\operatorname{anc}\left(\tilde{g}_n(v_i)d_{v_i}\left(\{t_j\}_{j=1}^i\right)\right)\right)\\ &= \tilde{h}\left(\tilde{g}_n\left(v_i\right)\right)\\ &= v_i\end{aligned}$$

for $i = 2, 3, \ldots, n$. The result follows. □

Let $\beta : \mathcal{G}_n \to \mathcal{R}_n$ be the function defined by

$$\beta(r, V, E, <, l) := (r, V, E).$$

Informally we say that $\beta$ "forgets" the ordered labelling of a rooted ordered tree.

**Proposition 3.1.5.** Suppose that $\{t_i\}_{i=1}^n \in \mathcal{T}_{w,n}$ and $s = g\left(\{t_i\}_{i=1}^n\right)$. Let $G$ be a rooted ordered tree with a historical sequence $s$. Then,

$$\beta(G) \cong t_n$$

*Proof.* Let $l_n : V(t_n) \to [n]$ be the labelling given in Definition 1.2.4 and suppose $V(t_n) = \{w_1, w_2, \ldots, w_n\}$ where $l_n(w_i) = i$ for all $i \in [n]$. Let $\mathbf{v}\left(\{t_i\}_{i=1}^n\right) = (v_2, v_3, \ldots, v_n)$ be the associated vertex sequence.

Suppose $s = (s_1, s_2, \ldots, s_n)$ and $l : V(G) \to \mathcal{N}$ is the labelling of $G$ given in Definition 2.2.2. Let $V(G) = \{x_1, \ldots, x_n\}$ and $s_i = l(x_i)$ for all $i \in [n]$.

Suppose $a : V(t_n) \to V(G)$ is the map defined by $a(w_i) = x_i$ for all $i \in [n]$. We claim that $a$ is a rooted isomorphism. Note that $a(w_1) = x_1 = \emptyset$ as required.

To prove this claim let $(w_i, w_j) \in V(t_n)$ and without loss of generality $i < j$. Then $w_i$ is the parent of $w_j$. Hence,

$$a(w_i) = x_{v_j}.$$

Since $s = g(\{t_i\}_{i=1}^n)$,

$$s_j = \tilde{g}_n(v_j) d_{v_j}\left(\{t_j\}_{j=1}^i\right) \text{ and } s_i = \tilde{g}_n(v_j)$$

This is the case if and only if $(x_i, x_j) \in V(\beta(G))$. □

**Corollary 3.1.6** (to Theorem 3.1.4)**.** Fix a rooted tree $t \in \mathcal{R}_n$

$$\sum_{G=\beta^{-1}(t)} \sum_{s \in S(G)} 1 = K(t).$$

Corollary 3.1.6 says that for a given rooted tree $t \in \mathcal{R}_n$, the tree multiplicity, $K(t)$, is the number of possible historical orderings of rooted ordered trees that, when we "forget" the ordering (under map $\beta$), correspond to $t$.

## 3.2 Network motifs

Commonly occurring subgraphs of real-world networks such as $(n, k)$-stars are known as *network motifs* and have been described as the building blocks of many real-world networks [MSOI+02].

In Section 2.1 we discovered that induced subgraphs contribute to the automorphism group of the whole rooted tree via the rooted geometric decomposition. In this section we will calculate the almost sure value of the contribution to the automorphism group of a particular subtree and prove Theorem 1.

Suppose $\mathcal{Y}_w^d$ is the discrete time Markov chain described in Section 2.3 and let $G_n = \mathcal{Y}_w^d(n)$. Subsequently, at time $n+1$, a new vertex $x_n k_n$ is attached to $x_n \in V(G_n)$ (here $k_n = \text{out}(x_n) + 1$) to build $G_{n+1} = G_n \cup \{x_n k_n\}$. In particular $s = (s_1, s_2, \ldots, s_n)$ is a historical ordering of $G_n$ and $s' = (s'_1, s'_2, \ldots, s'_n, s'_{n+1})$ given by $s_1 = s'_1 = \emptyset$, $s_i = s'_i = x_{i-1} k_{i-1}$ for all $i \in [n]$ and $s'_{n+1} = \{x_n k_n\}$ is a historical ordering of $G_{n+1}$. We remark that,

$$\mathbb{P}(\mathcal{Y}_w^d(n+1) = G_n \cup \{x_n k_n\}) = \mathbb{P}(s'_{n+1} = x_n k_n)$$

which is just the probability that vertex $x_n \in V(G_n)$ is chosen at time $n+1$.

Initially $\beta\left(\mathcal{Y}_w^d\right) = \bullet$. Let $t_n = \beta(G_n)$ and $t_{n+1} = \beta(G_{n+1})$. Suppose $h(s) = (v_2, v_3, \ldots, v_n)$, then $h(s') = (v_2, v_3, \ldots, v_{n+1}, \tilde{h}(x_n))$. In particular $t_{n+1}$ is built from $t_n$ by attaching a new vertex to $\tilde{h}(x_n)$ with probability

$$\mathbb{P}\left(s'_{n+1} = x_n k_n\right) = \frac{\widetilde{w}(\operatorname{out}(x_n, G_n))}{\sum_{y \in G_n} \widetilde{w}\left(\operatorname{out}(y, G_n)\right)}$$

Hence $\mathbb{P}(\beta\left(\mathcal{Y}_w^d(n+1)\right) = t) = \mathbb{P}(t_{n+1} = t : \{t_i\}_{i \in N} \in \mathcal{T}_w)$ and we say that $\mathcal{Y}_w^d$ is a *model for random attachment trees.*

Suppose $\mathcal{Y}_w$ is the continuous-time Markov process defined in Section 2.3. The jump chain $(Y_{w,n})_{n \in \mathbb{N}_0}$ associated with $\mathcal{Y}_w$ is a discrete time Markov chain distributed like $\mathcal{Y}_w^d$.

Suppose that $\{t_n\}_{n \in \mathbb{N}} \in \mathcal{T}_w$. Let $X_{n,k}$ be the number of vertices $v \in V(t_n)$ such that $\operatorname{out}(v, t_n) = k$ for $k \in \mathbb{N}$. Fix some rooted tree $t \in \mathcal{R}$. Let $Z_{n,t}$ be the number of induced subtrees of $\{t_n\}_{n \in \mathbb{N}}$ isomorphic to $t$.

**Proposition 3.2.1.** Fix $t \in \mathcal{R}$. Then,

$$\frac{X_{n,k}}{n} \xrightarrow{\text{a.s.}} p_w(k) \tag{3.2.1}$$

$$\frac{Z_{n,t}}{n} \xrightarrow{\text{a.s.}} \sum_{G = \beta^{-1}(t)} \pi_w(G) \tag{3.2.2}$$

for all $k \geq 0$.

In the next example we put together Proposition 3.2.1 with Example 2.3.4.

**Example 3.2.2.** Suppose $\{t_i\}_{i \in \mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$. Then,

(i)
$$\frac{X_{n,k}}{n} \xrightarrow{\text{a.s.}} 2^{-(k+1)}$$
for all $k \geq 0$. This is Janson's Theorem (Theorem 1.2.6). Hence we can think of Theorem 2.3.3 Part (i) as a generalisation of Janson's Theorem. This demonstrates the power of the continuous random tree method.

(ii) Fix $t \in \mathcal{R}$. Then,
$$\begin{aligned} \frac{Z_{n,t}}{n} &\xrightarrow{\text{a.s.}} \sum_{G = \beta^{-1}(t)} \sum_{s \in S(G)} \frac{1}{(|G| + 1)!} \\ &= \frac{K(t)}{(|t| + 1)!} \end{aligned}$$
by Corollary 3.1.6.

**Theorem 3.2.3** (continuous mapping theorem [Bil13])**.** Let $X \in \mathbb{R}$ and suppose that $\{X_n\}$ is a sequence of real-valued random variables. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Then,

(i) $X_n \xrightarrow{\text{a.s.}} X \Rightarrow f(X_n) \xrightarrow{\text{a.s.}} f(X)$.

(ii) Given any finite collection $\{X_n^1\}, \{X_n^2\}, \dots, \{X_n^m\}$ of sequences of random variables that almost surely converge to $X^1, X^2, \dots, X^m$ respectively, then

$$\sum_{i=1}^{m} \{X_n^i\} \xrightarrow{\text{a.s.}} \sum_{i=1}^{m} X^i,$$

for every $m \in \mathbb{N}$.

We define the *tree factorial* by $\bullet! := 1$ and then recursively by,

$$t! := |t| \prod_{i=1}^{k} t_i!$$

where $t$ is a root vertex adjacent to a rooted forest $f = \bigsqcup_{i=1}^{k} t_i$. The notation $t!$ is taken from [But08]. In Figure 3.4 we calculate the tree factorial for some examples.

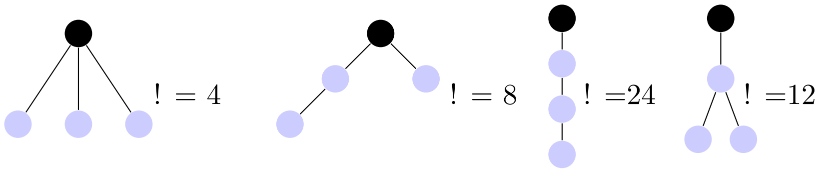


Figure 3.4: Examples of tree factorials

**Lemma 3.2.4.** [But08]. Suppose $t \in \mathcal{R}_n$. Then,

$$K(t) = \frac{|t|!}{t!\zeta(t)}. \tag{3.2.3}$$

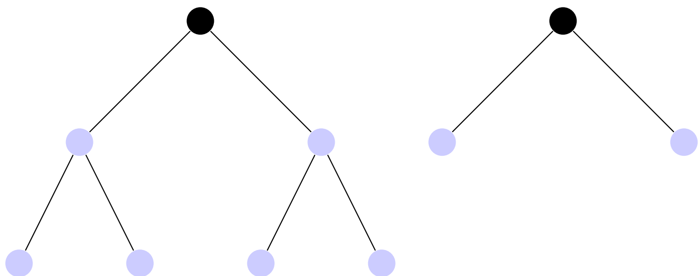

Figure 3.5: Two rooted trees $s \in \mathcal{R}_7$ and $s' \in \mathcal{R}_3$.

*Proof.* (Proof of Theorem 1) Suppose $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$ is a random recursive tree and fix the rooted trees $s \in \mathcal{R}_7$ and $s' \in \mathcal{R}_3$ shown in Figure 3.5. Let $Z_{s,n}$ be the number of induced subtrees of $t_n$ isomorphic to $s$.

$$\begin{aligned} K(s) &= \frac{7!}{63.8} \\ &= 10. \end{aligned}$$

By Example 3.2.2,

$$\left(\frac{Z_{s,n}}{n}\right) \xrightarrow{\text{a.s.}} \frac{K(s)}{(|s|+1)!} \tag{3.2.4}$$
$$= \frac{10}{8!} \tag{3.2.5}$$

Let $H(t_n) \leq \operatorname{Aut}(t_n)$ be the group that acts on every induced subtree of $t_n$ isomorphic to $s$ by swapping the two copies of $s'$. Clearly $|H(t_n)| = 2^{Z_{s,n}}$. By Theorem 3.2.3

$$\begin{aligned}\lim_{n\to\infty} \frac{\log\left(|H(t_n)|\right)}{n} &= \lim_{n\to\infty} \log(2)\frac{Z_{s,n}}{n} \\ &= \log(2)\frac{K(s)}{(|s|+1)!} \\ &= \frac{10\log(2)}{8!}\end{aligned}$$

almost surely. Since exponentiation is a continuous function, $|H(t_n)|^{\frac{1}{n}} \xrightarrow{\text{a.s.}} 2^{\frac{10}{8!}}$ Clearly $H(t_n) \leq \mathcal{C}(t_n)$ so $|H(t_n)| \leq |\mathcal{C}(t_n)|$ for all $n \in \mathbb{N}$.

All that is left to complete this proof is the rather obvious remark that $2^{\frac{10}{8!}} > 1$. $\square$

# Chapter 4

# Existence of limit

In [RTV07] Rudas, Tóth and Valkó show that a family of graphs called *rooted ordered trees* can be modelled as a family of continuous-time Markov chains called *Crump-Mode-Jagers (C-M-J)* processes. Rudas, Tóth and Valkó used results from the theory of C-M-J processes, notably those from [Ner81], to calculate local properties of rooted ordered trees such as degree distribution and the abundance of particular subtrees in terms of behaviour at the root vertex. We apply these theorems to random recursive trees.

In this chapter we give a rudimentary introduction to branching processes but the interested reader should see [Ner81] for more details on C-M-J general branching processes. We end this chapter with the proof of Theorem 2.

In Section 4.1 we define supercritical Malthusian branching processes following [Ner81]. We give the necessary background to and state Theorem 4.1.2 (Theorem 3.1 in [Ner81]) which gives the ratio of supercritical Malthusian branching processes counted by two characteristics. In Section 4.2 we use an elementary application of Theorem 4.1.2 to prove there exists $L \in \mathbb{R}$ such that $\zeta(t_n)^{\frac{1}{n}} \xrightarrow{\text{a.s.}} L$ (see Lemma 4.2.2). We end this section by proving $L = \mathcal{W}$ (see Theorem 2).

## 4.1 Branching processes

A *random branching process* consists of points that we call *individuals*. We call the collection of these individuals the *population*. Initially the population consists of a single individual called the *original individual*. The makeup of the population changes over time. Each individual reproduces like a random point process $\xi$.

We label the individuals in the population by $\mathcal{N}$ (see Definition 2.2.1). For example we write $x = (x_1, \ldots, x_n) \in \mathcal{N}$ for the individual that is the $x_n^{th}$ child of the $x_{n-1}^{th}$ child of the ... of the $x_1^{th}$ child of the original individual. The basic probability space for a C-M-J branching process is the product space,

$$(\Omega, \mathcal{B}, \mathbb{P}) = \prod_{x \in \mathcal{N}} (\Omega_x, \mathcal{B}_x, \mathbb{P}_x)$$

where all of the $(\Omega_x, \mathcal{B}_x, \mathbb{P}_x)$ are identical. On each space $(\Omega_x, \mathcal{B}_x, \mathbb{P}_x)$ we define point processes $\xi_x$ that are distributed like $\xi$.

Define $\nu := \mathbb{E}(\xi)$. We say that $\nu$ is the *intensity measure* of $\xi$ and we write,

$$\nu(\tau) := \mathbb{E}\left(\xi(\tau)\right).$$

We say that $\nu(\tau)$ is the *reproduction function* and we make the following assumptions:

(C1) The reproduction function $\nu(\tau)$ is not (as a measure) concentrated on any lattice $\{0, h, 2h, \dots\}$. For further details see [Ner81].

(C2) There exists a Malthusian parameter $\mathscr{m} \in (0, \infty)$, i.e. a finite positive solution of the equation:

$$\int_0^\infty e^{-\mathscr{m}\tau} \nu(d\tau) = 1$$

(C3) The first moment of $e^{-\mathscr{m}t}\nu(dt)$ is finite, i.e.

$$\int_0^\infty u e^{-\mathscr{m}u} \nu(du) < \infty.$$

A processes that satisfies conditions (C2) and (C3) is called a *supercritical Malthusian process* [Ner81].

Let $\nu_{\mathscr{m}}$ be the measure on $[0, \infty)$ defined by,

$$\nu_{\mathscr{m}}(\tau) = \int_0^\tau e^{-\mathscr{m}s} \nu(ds) \tag{4.1.1}$$

where $\mathscr{m}$ is the Malthusian parameter associated with the reproduction function.

Suppose that in addition to a point process $\xi$ there exists a product-measurable, separable, non-negative random process $\phi : \Omega \times \mathbb{R} \to \mathbb{R}$. Loosely speaking $\phi(\tau)$ assigns a value or score to a typical individual at age $\tau$. For simplicity we assume that

$$\phi(\tau) = 0 \text{ if } \tau < 0.$$

Let $\sigma_x$ be the birth time of the individual $x$. Following Nerman [Ner81], define

$$Z_\tau^\phi := \sum_{x \in \mathcal{N}} \phi(\tau - \sigma_x)$$

and we say that $\{Z_\tau^\phi\}_{\tau\in\mathbb{R}_\geq}$ is a *general branching process with characteristic* $\phi$.

**Example 4.1.1.** (i)

Define

$$\psi(\tau) := \begin{cases} 1 & \text{if } 0 \leq \tau < a \text{ and} \\ 0 & \text{otherwise} \end{cases} \tag{4.1.2}$$

then $Z_\tau^\psi$ counts the number of individuals alive at time $\tau$ whose ages are less than $a$.

(ii) Define,

$$\chi(\tau) := \begin{cases} 1 & \text{if } \tau \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.1.3}$$

then $Z_\tau^\chi$ counts the number of individuals alive at time $\tau$.

Suppose that $\phi$ is a characteristic. We write,

$$m_\tau^\phi = \mathbb{E}\left(e^{-\mathscr{m}\tau} Z_\tau^\phi\right).$$

We add two further assumptions. The first is an assumption on the random point process $\xi$ and the second is an assumption on the random process $\phi(\tau)$.

(C4) Suppose $\xi$ satisfies condition $(C1)-(C3)$. Then there exists $\mathscr{n} < \mathscr{m}$ such that $\nu_{\mathscr{n}}(\infty) < \infty$.

(C5) Suppose $\phi : \Omega \times \mathbb{R} \to \mathbb{R}$ is a product-measurable, separable, non-negative random process. There exists $\mathscr{n} < \mathscr{m}$ such that

$$V = \sup_{\tau\in\mathbb{R}} e^{-\mathscr{n}\tau}\phi(\tau)$$

has finite expectation.

**Theorem 4.1.2.** [Ner81, Theorem 3.1] Consider a supercritical Malthusian branching process $\xi$ which satisfies conditions (C1)-(C4), counted by two random characteristics $\phi_1$ and $\phi_2$ that satisfy condition (C5). Then, on $\{Z_\tau^\psi \to \infty\}$

$$\lim_{\tau\to\infty} \frac{Z_\tau^{\phi_1}}{Z_\tau^{\phi_2}} = \frac{m_\infty^{\phi_1}}{m_\infty^{\phi_2}}$$

almost surely.

Rudas, Tóth and Valkó prove that the general branching process $\xi_w$ that corresponds to the continuous-time model of an attachment tree satisfies conditions (C1) - (C4) [RTV07].

## 4.2 An application of Theorem 2.3.3

Rudas, Tóth and Valkó [RTV07] remark that the continuous-time random tree $\mathcal{Y}_w(\tau)$ has the same distribution as the time evolution of the continuous-time branching process with point process $\xi_w$ (given in Section 2.3.1). In particular, that general branching satisfies conditions (C1) - (C4). Furthermore the Malthusian parameter associated with the point process $\xi_w$ is $\lambda_w^*$ [RTV07]. For example, the Malthusian parameter associated with a random recursive tree is $\mathscr{m} = 1$.

Suppose $\xi_w$ is the continuous point process that corresponds to the continuous-time attachment tree $\mathcal{Y}_w(\tau)$. Let $\chi$ be as in Equation 4.1.3 respectively. Then,

$$\{Z_\tau^\chi\}_{\tau \in \mathbb{R}_\geq} = |\mathcal{Y}_w(\tau)|$$

is a general branching processes with characteristic $\chi$. For a given random characteristic $\phi$, [Ner81]

$$m_\infty^\psi = \int_0^\infty e^{-\tau} \mathbb{E}(\psi(\mathcal{Y}_w(\tau))) d\tau.$$

Define a new characteristic,

$$\theta(\tau) := \begin{cases} \log\left(\Lambda(\beta\left(\mathcal{Y}_{\widetilde{w}}(\tau)\right))\right) & \text{if } \tau \geq 0 \text{ and} \\ 0 & \text{otherwise.} \end{cases} \tag{4.2.1}$$

**Lemma 4.2.1.** Let $\theta : \Omega \times \mathbb{R} \to \mathbb{R}$ be the random process defined in Equation 4.2.1 and $\xi_{\widetilde{w}}$ be the point process that corresponds to $\mathcal{Y}_{\widetilde{w}}(\tau)$. There exists $\mathscr{n} < 1$ such that

$$V = \sup_{\tau \in \mathbb{R}} e^{-\mathscr{n}\tau} \theta(\tau)$$

has finite expectation.

*Proof.* We define a new random characteristic,

$$\tilde{\theta}(\tau) := \begin{cases} \log\left(\text{out}(\emptyset, \mathcal{Y}_{\widetilde{w}}(\tau))!\right) & \text{if } \tau \geq 0 \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Since $\theta(\tau) \leq \tilde{\theta}(\tau)$ for all $\tau \in \mathbb{R}$ it is enough to prove that

$$W := \sup_{\tau \in \mathbb{R}} e^{-\mathscr{n}\tau} \tilde{\theta}(\tau)$$

has finite expectation.

The holding times $H_1, H_2, \ldots,$ of $X_{\widetilde{w}}$ are exponentially distributed random variables with parameter $\lambda = 1$.

We remark that since $W = 0$ if $\tau < 0$ we need only consider the case $\tau \geq 0$. In particular $e^{-\varkappa\tau}$ is a monotonically decreasing function for $\tau \geq 0$ and $0 < \varkappa \leq 1$. Therefore,

$$\sup_{\tau\in\mathbb{R}} e^{-\varkappa\tau}\tilde{\theta}(\tau)$$

occurs at jump time $J_k$ for some $k \in \mathbb{N} \cup \{\infty\}$. Let $a_k := \tilde{\theta}(J_k)$ for all $k \in \mathbb{N}$. We have shown that,

$$\sup_{\tau\in\mathbb{R}} e^{-\varkappa\tau}\tilde{\theta}(\tau) = \sup_{k\in\mathbb{N}} e^{-\varkappa J_k} a_k < \infty,$$

for some $k \in \mathbb{N} \cup \{\infty\}$.

$\square$

**Lemma 4.2.2.** Let $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$. There exists a limit $L \in (0, \infty)$ such that

$$\zeta(t_n)^{\frac{1}{n}} \xrightarrow{\text{a.s.}} L.$$

*Proof.* Let $\theta$ be the random process defined in Equation 4.2.1 and $\chi$ be the random process defined in Equation 4.1.3.

Rudas, Tóth and Valkó prove that $\chi$ satisfies condition (C5) [RTV07]. By Lemma 4.2.1 $\theta$ satisfies condition (C5). Therefore, by Theorem 4.1.2 and Remark 2.1.1,

$$\int_0^\infty e^{-\tau}\mathbb{E}(\theta(\mathcal{Y}_{\widetilde{w}}(t)))dt = \lim_{\tau\to\infty} \frac{1}{|\mathcal{Y}_{\widetilde{w}}(\tau)|} \sum_{v\in\mathcal{Y}_{\widetilde{w}}(\tau)} \theta(\mathcal{Y}_{\widetilde{w}}(\tau)_{\downarrow v}) \tag{4.2.2}$$

$$= \lim_{\tau\to\infty} \frac{1}{|\mathcal{Y}_{\widetilde{w}}(\tau)|} \sum_{v\in\mathcal{Y}_{\widetilde{w}}(\tau)} \log(\Lambda(\mathcal{Y}_{\widetilde{w}}(\tau)_{\downarrow v}) \tag{4.2.3}$$

$$= \lim_{\tau\to\infty} \frac{\log\left(\zeta(\mathcal{Y}_{\widetilde{w}}(\tau))\right)}{|\mathcal{Y}_{\widetilde{w}}(\tau)|}. \tag{4.2.4}$$

almost surely. Since $\mathcal{Y}_{\widetilde{w}}$ is a model for random recursive trees,

$$\log\left(\zeta(t_n)^{\frac{1}{n}}\right) \xrightarrow{\text{a.s.}} \int_0^\infty e^{-t}\mathbb{E}(\theta(\mathcal{Y}_{\widetilde{w}}(\tau)))dt < \infty.$$

$\square$

Given a rooted tree $s \in \mathcal{R}$ and a rooted ordered tree $G$ define the indicator function

$$\widehat{I}^s(G) = \begin{cases} 1 & \text{if } s \cong \phi^{-1}(G) \\ 0 & \text{otherwise.} \end{cases}$$

The proof of Theorem 2 relies on the decomposition of $\zeta(t)$ in terms of the root permutation function given in Remark 2.1.1.

*Proof of Theorem 2.* By Remark 2.1.1,

$$\frac{\log(\zeta(\mathcal{Y}_{\widetilde{w}}(\tau)))}{|\mathcal{Y}_{\widetilde{w}}(\tau)|} = \frac{1}{|\mathcal{Y}_{\widetilde{w}}(\tau)|} \sum_{s \in \mathcal{R}} \sum_{v \in \mathcal{Y}_{\widetilde{w}}(\tau)} \log(\Lambda(s)) \widehat{I}^s(\mathcal{Y}_{\widetilde{w}}(\tau)_{\downarrow v}) \tag{4.2.5}$$

for all $\tau \in \mathbb{R}$. In particular,

$$\lim_{\tau\to\infty} \frac{1}{|\mathcal{Y}_{\widetilde{w}}(\tau)|} \log(\zeta(\mathcal{Y}_{\widetilde{w}}(\tau))) = \lim_{\tau\to\infty} \frac{1}{|\mathcal{Y}_{\widetilde{w}}(\tau)|} \sum_{s \in \mathcal{R}} \sum_{v \in \mathcal{Y}_{\widetilde{w}}(\tau)} \log(\Lambda(s)) \widehat{I}^s(\mathcal{Y}_{\widetilde{w}}(\tau)_{\downarrow v}).$$

Hence,

$$\lim_{\tau\to\infty} \frac{1}{|\mathcal{Y}_{\widetilde{w}}(\tau)|} \log(\zeta(\mathcal{Y}_{\widetilde{w}}(\tau))) = \lim_{\tau\to\infty} \sup_{S \subset \mathcal{R}, |S|<\infty} \frac{1}{|\mathcal{Y}_{\widetilde{w}}(\tau)|} \sum_{s \in S} \sum_{v \in \mathcal{Y}_{\widetilde{w}}(\tau)} \log(\Lambda(s)) \widehat{I}^s(\mathcal{Y}_{\widetilde{w}}(\tau)_{\downarrow v}).$$

We remark that for a fixed $s \in \mathcal{R}$

$$\lim_{\tau\to\infty} \frac{1}{|\mathcal{Y}_{\widetilde{w}}(\tau))|} \sum_{v \in \mathcal{Y}_{\widetilde{w}}(\tau)} \log(\Lambda(s)) \widehat{I}^s(\mathcal{Y}_{\widetilde{w}}(\tau)_{\downarrow v}) = \frac{K(s)\log(\Lambda(s))}{(|s|+1)!}$$

almost surely. Hence, by Theorem 3.2.3,

$$\begin{aligned}\lim_{\tau\to\infty} \frac{1}{|\mathcal{Y}_{\widetilde{w}}(\tau)|} \sup_{S \subset \mathcal{R}, |S|<\infty} \sum_{s \in S} \sum_{v \in \mathcal{Y}_{\widetilde{w}}(\tau)} \log(\Lambda(s)) \widehat{I}^s(\mathcal{Y}_{\widetilde{w}}(\tau)_{\downarrow v}) &= \sup_{S \subset \mathcal{R}, |S|<\infty} \sum_{s \in S} \frac{K(s)\log(\Lambda(s))}{(|s|+1)!} \\ &= \sum_{s \in \mathcal{R}} \frac{K(s)\log(\Lambda(s))}{(|s|+1)!}\end{aligned}$$

almost surely. Putting this together with Equation 4.2.5 we see that,

$$\lim_{\tau\to\infty} \frac{\log(\zeta(\mathcal{Y}_{\widetilde{w}}(\tau)))}{|\mathcal{Y}_{\widetilde{w}}(\tau)|} = \sum_{s \in \mathcal{R}} \frac{K(s)\log(\Lambda(s))}{(|s|+1)!}$$

almost surely. Hence, by Theorem 3.2.3,

$$\lim_{\tau\to\infty} \zeta(\mathcal{Y}_{\widetilde{w}}(\tau))^{\frac{1}{|\mathcal{Y}_{\widetilde{w}}(\tau)|}} = \exp\left(\sum_{s \in \mathcal{R}} \frac{K(s)\log(\Lambda(s))}{(|s|+1)!}\right)$$

The result follows since $\mathcal{Y}_{\widetilde{w}}(\tau)$ is a model for a random recursive tree. $\square$

We remark that,

$$\sum_{k=1}^{5} \sum_{t \in \mathcal{R}_k} \frac{K(t)\log(\Lambda(t))}{(|t|+1)!} = 0 + 0 + \frac{\log(2)}{4!} + \frac{\log(6)}{5!} + \frac{9\log(2)+\log(24)}{6!}.$$

Exponentiating we see that

$$\begin{aligned}
\mathcal{W} &\geq \exp\left(\frac{\log(2)}{4!} + \frac{\log(6)}{5!} + \frac{9\log(2) + \log(24)}{6!}\right) \\
&= \prod 2^{\frac{1}{4!}} 6^{\frac{1}{5!}} 2^{\frac{9}{6!}} 24^{\frac{1}{6!}} \\
&\approx 1.05854
\end{aligned}$$

In this way it is possible to get a numerical estimate for the limiting order of the automorphism group of a random recursive tree. We remark that since $1.05854... < \mathcal{V}$ we have not proved Theorem 3.

# Chapter 5

# Automorphisms of random recursive trees

In Chapter 4 we proved that, given $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$, $\zeta(t_n) \xrightarrow{\text{a.s.}} \mathcal{W}$. In this chapter we will achieve our long term goal of proving $\mathcal{W} \neq \mathcal{V}$, hence the intriguing relationship that MacArthur and Anderson identify in [MA06] is simply a coincidence.

We denote the subset of permutations that act on $V(t)$ by permuting copies of $s$, $\mathrm{Aut}^s(t)$. and we write $\zeta^s(t) := |\mathrm{Aut}^s(t)|$. Suppose $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$. In Section 5.4 and 5.5 respectively we show that there exist constants $\mathcal{L}_1$ and $\mathcal{L}_2$ respectively such that

$$\lim_{n\to\infty} \zeta^{\bullet}(t_n)^{\frac{1}{n}} > \mathcal{L}_1 \text{ and } \lim_{n\to\infty} \prod_{s\in\mathcal{R}\setminus\bullet} \zeta^s(t_n)^{\frac{1}{n}} > \mathcal{L}_2$$

with probability 1. For the proof of Theorem 3 we simply note that $\mathcal{V} < \mathcal{L}_1\mathcal{L}_2 < \mathcal{W}$.

In Section 5.2 and 5.5 we define a standard combinatorial map $\widehat{T} : S_n \to \mathcal{T}_{w,n+1}$ and a novel map $\tilde{j} : S_n \to \mathbb{R}$ respectively, such that the following triangle commutes:

$$\begin{tikzcd} & S_n \arrow[dl, "\widehat{T}"'] \arrow[d, "\tilde{j}"] \\ \mathcal{T}_{w,n+1} \arrow[r, "\log(\Lambda)"'] & \mathbb{R}. \end{tikzcd} \tag{5.0.1}$$

In Section 5.3 we use the algebraic machinery of the *cycle indicator polynomial* to express $\lim_{n\to\infty} \frac{\log(\zeta^{\bullet}(t_n))}{n}$ in terms of an integral (see Proposition 5.3.7). In Section 5.4 we calculate this integral using a novel family of polynomials called *recursively exponential polynomials* which allows us to calculate $\mathcal{L}_1$. In Section 5.5 we calculate $\mathcal{L}_2$ and we end this section by proving Theorem 3.

Before be begin the chapter proper we introduce further notation.

**Notation.** Throughout this chapter we fix a rooted tree $s \in \mathcal{R}$.

Fix $k \in \mathbb{N}$. We define,

$$\mathcal{R}_n^{s,k} := \{t \in \mathcal{R}_n : t = B^+(s^k, t_2^{\alpha_2}, \dots, t_m^{\alpha_m})\}$$

and $\mathcal{R}^{s,k} := \bigcup_{n\in\mathbb{N}} \mathcal{R}^{s,k}$. Define an indicator function,

$$I^{s,k}(t) := \begin{cases} k & \text{if } t = B^+(s^k, t_2^{\alpha_2}, \dots, t_m^{\alpha_m}) \\ 1 & \text{otherwise.} \end{cases}$$

for all $k \geq 1$. Define a further indicator function

$$\widetilde{I}^s(t) = \begin{cases} 1 & \text{if } t = B^+(s^{\alpha_1}, t_2^{\alpha_2}, \dots, t_m^{\alpha_m}) \\ 0 & \text{otherwise.} \end{cases}$$

## 5.1 Limiting contribution of $k$-stars

Let $\mathrm{Aut}^{s,k}(t) \leq \mathrm{Aut}^s(t)$ be the group that acts on $t$ by permuting copies of $s$ whenever $t_{\downarrow v} = B^+(s^k, t_2^{\alpha_2}, \dots, t_m^{\alpha_m})$. In particular,

$$\mathrm{Aut}^{s,k}(t) := \prod_{v\in V} S_{I^{s,k}(t_{\downarrow v})},$$

and we write $\zeta^{s,k}(t) := \left|\mathrm{Aut}^{s,k}(t)\right|$ for any rooted tree $t \in \mathcal{R}$.

**Lemma 5.1.1.** Let $\{t_i\}_{i\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$ and fix $k \in \mathbb{N}$. Then,

$$\frac{\log(\zeta^{\bullet,k}(t_n))}{n} \xrightarrow{\text{a.s.}} \sum_{t\in\mathcal{R}^{\bullet,k}} \frac{K(t)\log(k!)}{(|t|+1)!}. \tag{5.1.1}$$

*Proof.* We remark that for any rooted tree $t \in \mathcal{R}_n$,

$$\begin{aligned}
\frac{\log(\zeta^{\bullet,k}(t))}{n} &= \frac{\log\left(\left|\prod_{v\in V} S_{I^{\bullet,k}(t_{\downarrow v})}\right|\right)}{n} \\
&= \frac{\sum_{v\in V} \log\left|S_{I^{\bullet,k}(t_{\downarrow v})}\right|}{n} \\
&= \frac{\sum_{v\in V} \log\left(I^{\bullet,k}(t_{\downarrow v})!\right)}{n} \\
&= \frac{\sum_{v\in V(t_n)} \sum_{s\in\mathcal{R}^{\bullet,k}} \widetilde{I}^s(t_{\downarrow v})\log(k!)}{n}
\end{aligned}$$

for all $k \geq 1$. The result follows by Example 3.2.2. $\square$

In order to calculate $\sum_{t \in \mathcal{R}^{\bullet,k}} \frac{K(t)\log(k!)}{(|t|+1)!}$ in the next section we will exploit a bijection between $(n+1)$-saplings and the symmetric group $S_n$.

## 5.2 A correspondence between $(n+1)$-saplings and $S_n$

In this section we will construct three standard maps: $T$,$\widehat{T}$ and $\widetilde{T}$. For a more thorough investigation of these maps thge interested reader should see [Pit95].

Let $\widehat{T}(\sigma) = \{t_i\}_{i=1}^{n+1}$. In Proposition 5.2.3 and Proposition 5.2.4 we will prove that $\sigma$ has *cycle type* $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ if and only if $t_{n+1}$ consists of a root vertex $r$ adjacent to a rooted forest $f = \bigsqcup_{k \in K} s_k$ of rooted trees such that

$$a_i = |\{k \in K : |s_k| = i\}| \,. \tag{5.2.1}$$

for all $i \in [n]$. Define,

$$j(\sigma) := \sum_{i=1}^{n} \log(a_i!).$$

In Proposition 5.2.5 we prove that $j(\sigma)$ is an upper bound for $\log(\Lambda(\widehat{T}(\sigma)))$.

In addition to permuting the vertices of a tree, the symmetric group has a deeper connection with random recursive trees hinted at by the observation

$$|\mathcal{T}_{w,n+1}| = |S_n| = n!$$

We begin this section by constructing an explicit bijection between $\mathcal{T}_{w,n+1}$ and $S_n$ via a class of nested sequences of permutations called *consistent random permutations* for each $n \in \mathbb{N}$.

Consider a sequence of permutations, $\{\sigma_m\}_{m=1}^{n}$ such that:

(i) each permutation $\sigma_m \in S_m$ and

(ii) if $\sigma_m$ is written as a product of cycles then $\sigma_{m-1}$ is derived by the deletion of element $m$ from the cycle of which it is a part. For example, if $\sigma_7 = (1)(542)(673)$ then $\sigma_6 = (1)(542)(63)$. If $\sigma_7 = (7)(542)(613)$ then $\sigma_6 = (542)(613)$.

We call these sequences *consistent random permutations*. As we remarked in the introduction to this sectThe set of all consistent permutations of length $n$ is denoted $\Sigma_n$.

**Definition 5.2.1.** [Pit95]

(i) There is a bijection $\widetilde{T} : \Sigma_n \longrightarrow S_n$ given by $\widetilde{T}(\{\sigma_i\}_{i=1}^n) = \sigma_n$.

(ii) There is a bijection $T : \Sigma_n \to \mathcal{T}_{w,n+1}$. Explicitly we have $T(\{\sigma_i\}_{i=1}^n) = (v_2, v_3 \ldots, v_{n+1})$ where $v_2 = 1$ and subsequently,

$$v_{i+1} = \begin{cases} 1 & \text{if } \sigma_i(i) = i \\ \sigma_i(i) + 1 & \text{otherwise} \end{cases} \tag{5.2.2}$$

for $i = 2, 3, \ldots, n$.

Suppose $\{t_i\}_{i=1}^{n+1} \in \mathcal{T}_{w,n+1}$ and let $l_{n+1} : V(t_{n+1}) \to [n+1]$ be the labelling of $t_{n+1}$ given in Section 3.1. Define a new labelling $l'_{n+1} : V(t_{n+1}) \to \{0, 1, \ldots, n\}$ defined by $l'_{n+1}(v) := l_n(v) - 1$ for any vertex $v \in V(t_{n+1})$.

Suppose the vertex sequence $\mathbf{v}(\{t_i\}_{i=1}^{n+1}) = (v_2, v_3, \ldots, v_{n+1})$ so that the label of the vertex $u$ attached to at time $i \in \{2, 3, \ldots, n+1\}$ is $l_{n+1}(u) = v_i$. We define the *lower vertex sequence*,

$$\mathbf{v}'(\{t_i\}_{i=1}^{n+1}) := (y_1, y_2, \ldots, y_n).$$

where $y_{i-1} = l'_i(v_i)$ for $i = 2, 3, \ldots, n+1$. In particular $y_{i-1} = v_i - 1$ for $i = 2, 3, \ldots, n+1$. Suppose vertex $v \in V(t_{n+1})$ and $l'_{n+1}(v) = i$ for some $i \in [n-1]$. The parent $w \in V(t_{n+1})$ of $v$ is labelled $l'_{n+1}(w) = y_i$.

**Definition 5.2.2.** [Pit95] The inverse map $T^{-1} : \mathcal{T}_{w,n+1} \to \Sigma_n$ is defined by $\sigma_1(1) = 1$, then recursively for $i < j$ by

$$\sigma_j(i) = \begin{cases} j & \text{if } \sigma_{j-1}(i) = y_j \\ \sigma_{j-1}(i) & \text{otherwise, and} \end{cases}$$

$$\sigma_j(j) = \begin{cases} j & \text{if } y_j = 0 \\ y_j & \text{otherwise} \end{cases}$$

for $j = 2, 3, \ldots, n$.

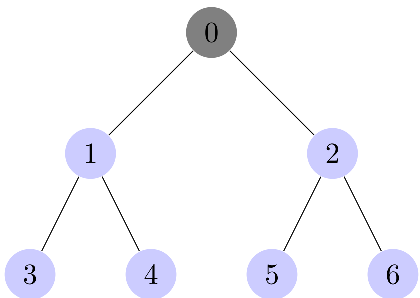


Figure 5.1: The random recursive tree $T((1), (1)(2), (3,1)(2), (3,4,1)(2), (3,4,1)(5,2), (3,4,1)(5,6,2))$.

Define a third bijection $\widehat{T} : S_n \to \mathcal{T}_{w,n+1}$ by,

$$\widehat{T}(\sigma) := T(\widetilde{T}^{-1}(\sigma))$$

for all $\sigma \in S_n$.

If it is not immediately obvious which permutation we are talking about then we write $a_i(\sigma)$ for the number of $i$-cycles in $\sigma$. We write $S_n^{\mathbf{a}}$ for the set of permutations $\sigma \in S_n$ such that $\sigma$ has cycle type $\mathbf{a}$ .

**Proposition 5.2.3.** Let $\sigma \in S_n$ and $\widehat{T}(\sigma) = \{t_i\}_{i=1}^{n+1}$. Suppose $t_{n+1}$ consists of a root vertex $r$ adjacent to a rooted forest $f = \bigsqcup_{k \in K} s_k$ of rooted trees. Then $\sigma$ has cycle type $(a_1, a_2, \ldots, a_n)$ where each $a_i$ is given by Equation 5.2.1.

*Proof.* Let the lower vertex sequence, $\mathbf{v}'\left(\{t_i\}_{i=1}^{n+1}\right) = (y_1, y_2, \ldots, y_n)$.

Fix $k \in K$ and suppose $V(s_k) = \{x_1, x_2, \ldots, x_p\}$ for some $p \in [n]$. Throughout this proof we will abuse notation and write $x_i = l'_{n+1}(x_i)$ for all $i \in [p]$. Without loss of generality suppose that $x_1 < x_2 < \cdots < x_p$. Since $s_k \in \mathcal{R}$ there is a root vertex $r(s_k)$. Clearly $x_1 = r(s_k)$ and the parent vertex of $x_1$ is $r(t_{n+1})$. Since $l'_{n+1}(r(t_{n+1})) = 0$, $y_{x_1} = 0$. The rooted tree $s_k$ is connected so $y_{x_i} \in V(s_k)$ for $i = 2, 3, \ldots, p$. In particular $y_{x_i} > 0$ for $i = 2, 3, \ldots, p$.

Let $\{\sigma_i\}_{i=1}^n \in \Sigma_n$ be such that $\widetilde{T}\left(\{\sigma_i\}_{i=1}^n\right) = \sigma$. By Definition 5.2.2 $\sigma_{x_1}(x_1) = x_1$. Suppose $j \in [p-1]$ and $i \leq j$. Then, by Definition 5.2.2,

$$\sigma_{x_j}(x_i) = \sigma_{x_j+1}(x_i) = \cdots = \sigma_{x_{j+1}-1}(x_i).$$

Therefore it is enough to consider $\sigma_{x_j}(x_i)$ for $j \in \{2, 3, \ldots, p\}$ and $i \leq j$. We have $\sigma_{x_j}(x_j) = y_{x_j}$ and,

$$\sigma_{x_j}(x_i) = \begin{cases} x_j & \text{if } \sigma_{x_{j-1}}(x_i) = y_{x_j} \\ \sigma_{x_{j-1}}(x_i) & \text{otherwise,} \end{cases} \tag{5.2.3}$$

by Definition 5.2.2. Clearly then, $\sigma_{x_j}(x_i) \in \{x_1, x_2, \ldots, x_j\}$ for all $j \in [p]$ and $x_i \leq x_j$. We claim that $\{x_1, x_2, \ldots, x_j\}$ is a cycle in $\sigma_{x_j}$ for all $j \in [p]$.

We will prove the claim by induction. We have $\sigma_{x_1}(x_1) = x_1$ which is a cycle of length 1 in $\sigma_{x_j}$. Suppose, for the inductive hypothesis, that $(c_1, c_2, \ldots, c_j)$ is a cycle in $\sigma_{x_j}$ and that each $c_i \in \{x_1, x_2, \ldots, x_j\}$.

Note that $y_{x_{j+1}}$ is the label of the parent of vertex $x_{j+1}$ hence, $y_{x_{j+1}} \in \{x_1, \ldots, x_j\}$. Suppose, without loss of generality, that $y_{x_{j+1}} = c_1$. Then, by Equation 5.2.3, $(x_{j+1}, c_1, c_2, \ldots, c_j)$ is a cycle in $\sigma_{x_{j+1}}$. □

Proposition 5.2.4 is the converse to Proposition 5.2.3.

**Proposition 5.2.4.** Suppose $\sigma \in S_n^{\mathbf{a}}$ and $\widehat{T}(\sigma) = \{t_i\}_{i=1}^{n+1}$. Then $t_{n+1}$ consists of a root vertex $r$ adjacent to a rooted forest $f = \bigsqcup_{k \in K} s_k$ of rooted trees. Then $a_i$ satisfies Equation 5.2.1 for all $i \in [n]$.

*Proof.* Let $\mathbf{v}\left(\{t_i\}_{i=1}^{n+1}\right) = (v_2, \ldots, v_{n+1})$ and suppose $C = (c_1, c_2, \ldots, c_r)$ is an $r$-cycle in $\sigma$. Let $X = \{x_1, x_2, \ldots, x_r\}$ with each $x_i \in [n]$ ordered so that $x_1 < x_2 < \cdots < x_r$ and each $c_i \in X$ is distinct.

Suppose $\widetilde{T}(\{\sigma_i\}_{i=1}^{n}) = \sigma$ and $1 \leq p < r$. By the construction of a consistent permutation $C_p = (c_{i_1}, c_{i_2}, \ldots, c_{i_p})$ is a cycle in the permutations $\sigma_{x_p}, \sigma_{x_p+1}, \ldots, \sigma_{x_{p+1}-1}$. In particular $\sigma_{x_1}(x_1) = x_1$ so by Definition 5.2.1 $v_{x_1+1} = 1$.

Since $C_p$ is a cycle we have $v_{x_p+1} \in X$ for $p = 2, 3, \ldots, r$. Since cycles are disjoint, $v_{i+1} \in X$ if and only if $i \in X \backslash \{x_1\}$.

Since each $v_i$ is the parent of the vertex $w$, where $l_{n+1}(w) = i$ the result follows. $\square$

We have begun building the machinery required to define $\tilde{j} : S_n \to \mathbb{R}$ such that triangle 5.0.1 commutes.

**Proposition 5.2.5.** Fix $n \in \mathbb{N}$. Suppose $\sigma \in S_n$ has cycle type $(a_1, a_2, \ldots, a_n)$ and $\widehat{T}(\sigma) = \{t_i\}_{i=1}^{n+1}$. Then,

$$\log(\Lambda(t_{n+1})) \leq j(\sigma). \tag{5.2.4}$$

*Proof.* Suppose $t_{n+1} = B^+(s_1^{\alpha_1}, s_2^{\alpha_2}, \ldots, s_m^{\alpha_m})$ and

$$a_j = \sum_{k=1}^{m} \sum_{s_k \in \mathcal{R}_j} \alpha_j \tag{5.2.5}$$

for $j = 2, 3, \ldots, |t| - 1$. By Proposition 5.2.4 $\widehat{T}(\sigma)$ has cycle type $(a_1, a_2, \ldots, a_n)$. Therefore,

$$\begin{aligned} j(\sigma) &= \sum_{i=1}^{n} \log(a_j!) \\ &\geq \sum_{i=1}^{m} \log(\alpha_i!) \end{aligned}$$

since for all positive integers $p, q \in \mathbb{N}$ we have $p! + q! \leq (p+q)!$. $\square$

Under the hypotheses given in the statement of Proposition 5.2.5 $j(\sigma)$ is an upper bound for $\log(\Lambda(t_{n+1}))$.

Fix $n \in \mathbb{N}$ and $k \in [n]$. Define,

$$S_n^{a_1=k} := \{\sigma \in S_n : \sigma \text{ has cycle type } (k, a_2, \dots, a_n)\}.$$

**Proposition 5.2.6.** Fix $k \in \mathbb{N}$. Then,

$$\sum_{t \in \mathcal{R}_{n+1}^{\bullet,k}} \frac{K(t)\log(k!)}{(|t|+1)!} = \sum_{\sigma \in S_n^{a_1=k}} \frac{\log(k!)}{(n+2)!}. \tag{5.2.6}$$

*Proof.* Suppose $\sigma \in S_n$ and $\widehat{T}(\sigma) = \{t_i\}_{i=1}^{n+1}$. Then $t_{n+1}$ consists of a root vertex $r$ adjacent to a rooted forest $f = \bigsqcup_{i \in I} s_i$ of rooted trees. By Proposition 5.2.3 and Proposition 5.2.4 $a_1(\sigma) = k$ if and only if $k = |\{i \in I : s_i \cong \bullet\}|$. Hence

$$\left|\left\{\{t_i\}_{i=1}^{n+1} \in \mathcal{T}_{w,n+1} : t_{n+1} = B^+(\bullet^k, t_2^{\alpha_2}, \dots, t_m^{\alpha_m})\right\}\right| = \left|S_n^{a_1=k}\right|$$

and the result follows from Definition 1.2.8. $\square$

Fix $n \in \mathbb{N}$ and $i, k \in [n]$. Let $X_n^{i,k}$ be the number of permutations $\sigma \in S_n$ such that $a_i(\sigma) = k$. In other words $X_n^{i,k}$ is the number of permutations $\sigma \in S_n$ with precisely $k$ cycles of length $i$. Then, $\left|S_n^{a_1=k}\right| = X_n^{1,k}$ for all $n \in \mathbb{N}$ and $k \in [n]$.

**Example 5.2.7.** The first non-zero values of $X_n^{i,k}$ are as follows,

$$\begin{aligned} X_1^{1,1} &= 1 \\ X_2^{2,1} &= 1 \\ X_2^{1,2} &= 1 \\ X_3^{1,3} &= 1 \\ X_3^{3,1} &= 2 \\ X_3^{1,1} &= 3 \\ X_3^{2,1} &= 3 \end{aligned}$$

## 5.3 The cycle indicator polynomial

In this section we will build the algebraic machinery necessary to calculate the righthand side of Equation 5.2.6. In particular we will use a well-known algebraic formulation of the symmetric group called the *cycle indicator polynomial.*

We begin this section with the necessary definitions and background information. We state Theorem 5.3.4 (Theorem 1.3.3 in [Sta86]) in which lays the power of the cycle indicator polynomial. In addition we prove Proposition 5.3.5 which is a standard application of this theorem. In Section 5.3.1 we define a 1-parameter family of generating functions, $H_k(x)$, that we will use to calculate the limiting behaviour of $\zeta(t_n)^{\frac{1}{n}}$.

Suppose that a permutation $\sigma \in S_n$ has cycle type $(a_1, \ldots, a_n)$. Write

$$z^{\text{type}(\sigma)} := z_1^{a_1} z_2^{a_2} \ldots z_n^{a_n}$$

and define the *cycle indicator polynomial*

$$Z_n := \frac{1}{n!} \sum_{\sigma \in S_n} z^{\text{type}(\sigma)}.$$

We define $Z_0 := 1$.

**Example 5.3.1.**

$$\begin{aligned}
Z_0 &= 1 \\
Z_1 &= z_1 \\
Z_2 &= \frac{1}{2!}(z_1^2 + z_2) \\
Z_3 &= \frac{1}{3!}(z_1^3 + 3z_1 z_2 + 2z_3) \\
Z_4 &= \frac{1}{4!}(z_1^4 + 6z_1^2 z_2 + 8z_1 z_3 + 3z_2^2 + 6z_4).
\end{aligned}$$

Define,

$$\widetilde{S}_n := S_n / \sim$$

where two elements $\sigma, \tau \in S_n$ are equivalent by $\sim$ if they belong to the same conjugacy class (i.e. they have the same cycle type). Write $[\sigma] \in \widetilde{S}_n$ for the conjugacy class representative of $\sigma \in S_n$. If $\sigma$ has cycle type $(a_1, a_2, \ldots, a_n)$ then the conjugacy class $[\sigma]$ has order $|[\sigma]| = \prod_{i=1}^n \frac{n!}{a_i! i^{a_i}}$ [Sta86]. Define a family of monomials,

$$Z_n^{[\sigma]}(z_1, \ldots, z_n) := |[\sigma]| z_1^{a_1} \ldots z_n^{a_n}.$$

Then,

$$Z_n(z_1, \ldots, z_n) = \frac{1}{n!} \sum_{[\sigma] \in \widetilde{S}_n} Z_n^{[\sigma]}(z_1, \ldots, z_n). \tag{5.3.1}$$

**Proposition 5.3.2.** Suppose $\sigma \in S_n^{\mathbf{a}}$. Then,

$$\frac{\partial^k}{\partial z_1^k} Z_n^{[\sigma]}(z_1, \ldots, z_n) \Big|_{\substack{z_1 = 0 \\ z_j = 1 \text{ for } j > 1}} = \begin{cases} |[\sigma]|\, a_1! & \text{if } a_1 = k \\ 0 & \text{otherwise.} \end{cases}$$

*Proof.* We split this proof into 3 cases.

**Case 1** $(k < a_1)$**.** Then,

$$\frac{\partial^k}{\partial z_1^k} Z_n^{[\sigma]}(z_1, \ldots, z_n) = |[\sigma]| \frac{a_1!}{(a_1 - k)!} z_1^{a_1 - k} z_2^{a_2} \ldots z_n^{a_n}.$$

Hence

$$\begin{aligned} \frac{\partial^k}{\partial z_1^k} Z_n^{[\sigma]}(z_1, \ldots, z_n)\Big|_{\substack{z_1=0 \\ z_j=1 \text{ for } j>1}} &= |[\sigma]| \frac{a_1!}{(a_1 - k)!} 0^{a_1 - k} 1^{a_2} \ldots 1^{a_n} \\ &= 0. \end{aligned}$$

**Case 2** $(k > a_1)$**.** Then,

$$\frac{\partial^k}{\partial z_1^k} Z_n^{[\sigma]}(z_1, \ldots, z_n) = 0.$$

**Case 3** $(k = a_1)$**.** Then,

$$\frac{\partial^k}{\partial z_1^k} Z_n^{[\sigma]}(z_1, \ldots, z_n) = |[\sigma]| \, a_1! z_2^{a_2} \ldots z_n^{a_n},$$

hence

$$\begin{aligned} \frac{\partial^k}{\partial z_1^k} Z_n^{[\sigma]}(z_1, \ldots, z_n)\Big|_{\substack{z_1=0 \\ z_j=1 \text{ for } j>1}} &= |[\sigma]| a_1! 1^{a_2} \ldots 1^{a_n} \\ &= |[\sigma]| a_1! \end{aligned}$$

□

**Corollary 5.3.3.** Fix $k \in [n]$. Then,

$$\frac{\partial^k}{\partial z_1^k} Z_n(z_1, z_2, \ldots, z_n)\Big|_{\substack{z_1=0 \\ z_j=1 \text{ for } j>1}} = \frac{1}{n!} \sum_{\sigma \in S_n^{a_1 = k}} k! \tag{5.3.2}$$

*Proof.* Let $i, k \in \mathbb{N}$. By Equation 5.3.1,

$$\begin{aligned} \frac{\partial^k}{\partial z_1^k} Z_n(z_1, \ldots, z_n) &= \frac{\partial^k}{\partial z_1^k} \frac{1}{n!} \sum_{[\sigma] \in \widetilde{S}_n} Z_n^{[\sigma]}(z_1, \ldots, z_n) \\ &= \frac{1}{n!} \sum_{\sigma \in S_n^{a_1 = k}} k! \end{aligned}$$

by Proposition 5.3.2 and since the derivative is linear. □

The cycle indicator polynomial is a particularly elegant and useful formulation of the symmetric group because of the following theorem.

**Theorem 5.3.4.** (This is Theorem 1.3.3 in [Sta86].)

$$\sum_{n\geq 0} Z_n(z_1, z_2, \ldots, z_n)x^n = \exp\left(z_1 x + z_2\frac{x^2}{2} + z_3\frac{x^3}{3} + \cdots\right).$$

Using a standard application of Theorem 5.3.4 we can extend the statement of Corollary 5.3.3 as follows.

**Proposition 5.3.5.**

$$\frac{\partial^k}{\partial z_i^k}\sum_{n\geq 0} Z_n(z_1, \ldots, z_n)x^n\Bigg|_{\substack{z_1=0\\ z_j=1 \text{ for } j>1}} = \frac{x^{ki}}{i^k}\left(\frac{1}{1-x}\right)e^{-\left(\frac{x^i}{i}\right)}.$$

*Proof.* By Theorem 5.3.4,

$$\begin{aligned}\frac{\partial^k}{\partial z_i^k}\sum_{n\geq 0} Z_n(z_1, \ldots, z_n)x^n &= \frac{\partial^k}{\partial z_i^k}\exp\left(z_1 x + z_2\frac{x^2}{2} + z_3\frac{x^3}{3} + \cdots\right)\\ &= \frac{x^{ki}}{i^k}\exp\left(z_1 x + z_2\frac{x^2}{2} + z_3\frac{x^3}{3} + \cdots\right).\end{aligned}$$

Hence,

$$\begin{aligned}\frac{\partial^k}{\partial z_i^k}\sum_{n\geq 0} Z_n(z_1, \ldots, z_n)x^n\Bigg|_{\substack{z_1=0\\ z_j=1 \text{ for } j>1}} &= \frac{x^{ki}}{i^k}\exp\left(x + \cdots\frac{x^{i-1}}{i-1} + \frac{x^{i+1}}{i+1} + \cdots\right)\\ &= \frac{x^{ki}}{i^k}\exp\left(\log\left(\frac{1}{1-x}\right) - \frac{x^i}{i}\right)\\ &= \frac{x^{ki}}{i^k}\left(\frac{1}{1-x}\right)e^{-\left(\frac{x^i}{i}\right)}.\end{aligned}$$

□

### 5.3.1 A family of generating functions

The most important result in this section is Theorem in which we show that for a random recursive tree the limiting value of $\zeta(t_n)$ can be expressed in terms of the generating function $H_k(x)$.

Fix $n \in \mathbb{N}$ and $k \in [n]$. Define a family of constants,

$$c_{n,k} := \sum_{\sigma\in S_n^{a_1=k}} k!$$

If $k > n$ we define $c_{n,k} := 0$ and we define $c_{0,0} := 0$. We remark that $S_n^{a_1=k}$ is the set of permutations $\sigma \in S_n$ with $k$ fixed points. Further, $\left|S_n^{a_1=k}\right| = n!\sum_{j=k}^{n}\frac{(-1)^{j-k}}{k!(j-k)!}$ are know as the *recontres numbers*[Rio12]. Hence,

$$c_{n,k} = n! \sum_{j=k}^{n} \frac{(-1)^{j-k}}{(j-k)!}.$$

Define a 1-parameter family of exponential generating functions,

$$H_k(x) := \sum_{n \geq k+2} \frac{c_{n-2,k} x^n}{n!}.$$

Let $X := \sum_{j=k}^{n-2} \frac{(-1)^{j-k}}{(j-k)!}$. By the "ratio test";

$$\begin{aligned} \lim_{n\to\infty} \left| \frac{c_{n-1,k} x^{n+1}/(n+1)!}{c_{n-2,k} x^n/n!} \right| &= \lim_{n\to\infty} \left| \frac{x(n-1)\left(X + \frac{(-1)^{n-1-k}}{(n-1-k)!}\right)}{(n+1)X} \right| \\ &= |x|, \end{aligned}$$

hence the radius of convergence of $H_k(x)$, $\rho_{H,k} = 1$. To determine convergence in the case $x = 1$ we remark that,

$$\begin{aligned} H_k(1) &= \sum_{n \geq k+2} \frac{c_{n-2,k}}{n!} \\ &= \sum_{n \geq k+2} \frac{1}{n(n-1)} \sum_{j=k}^{n-2} \left( \frac{(-1)^{j-k}}{(j-k)!} \right). \end{aligned}$$

Since $0 \leq \sum_{j=k}^{n-2} \left( \frac{(-1)^{j-k}}{(j-k)!} \right) \leq 1$ for all $n \in \mathbb{N}$, $H_k(1)$ converges absolutely by the "direct comparison test".

We have now developed all of the necessary machinery to state the primary theorem of this section.

**Theorem 5.3.6.** Fix $k \in \mathbb{N}$ and suppose $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$. Then,

$$\frac{\log\left(\zeta^{\bullet,k}(t_n)\right)}{n} \xrightarrow{\text{a.s.}} \frac{\log(k!)}{k!} H_k(1).$$

*Proof.* By Lemma 5.1.1,

$$\frac{\log\left(\zeta^{\bullet,k}(t_n)\right)}{n} \xrightarrow{\text{a.s.}} \sum_{t' \in \mathcal{R}^{\bullet,k}} \frac{K(t') \log(k!)}{(|t'|+1)!}$$

By Proposition 5.2.6,

$$\begin{aligned}\sum_{t\in\mathcal{R}^{\bullet,k}} \frac{K(t)\log(k!)}{(n+2)!} &= \sum_{n\geq 1}\sum_{t\in\mathcal{R}^{\bullet,k}_{n+1}} \frac{K(t)\log(k!)}{(n+2)!} \\ &= \sum_{n\geq 1}\sum_{\sigma\in S_n^{a_1=k}} \frac{\log(k!)}{(n+2)!} \\ &= \frac{\log(k!)}{k!}\sum_{n\geq 1}\frac{c_{n,k}}{(n+2)!} \\ &= \frac{\log(k!)}{k!}H_k(1).\end{aligned}$$

□

Suppose $F(x) = \sum_{n\geq 0}\frac{a_n}{n!}x^n$ is a formal power series. The standard definition of the *formal derivative* $F'(x) := \sum_{n\geq 0}\frac{a_{n+1}}{n!}x^n$ can be found, for example, in [Sta86]. We define $G_k(x) := H_k'(x)$ and $F_k(x) := G_k'(x)$. By Corollary 5.3.3,

$$\begin{aligned}F_k(x) &= \sum_{n\geq 0}\frac{c_{n,k}x^k}{n!} \\ &= \sum_{n\geq 0}\frac{\partial^k}{\partial z_1^k}Z_n(z_1,z_2,\ldots,z_n)x^n\Bigg|_{\substack{z_1=0\\ z_j=1 \text{ for } j>1}}\end{aligned}$$

Since the radius of convergence of $\rho_{H,k}=1$, the radii of convergence of $F_k(x)$ and $G_k(x)$ satisfy $\rho_{F,k}=1$ and $\rho_{G,k}=1$ respectively. By Proposition 5.3.5, $F_k(x) = \frac{x^k e^{-x}}{(1-x)}$.

It is now that we unleash the awesome power of Theorem 5.3.4.

**Proposition 5.3.7.** Suppose $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\widetilde{w}}$. Then,

$$\frac{\log\left(\zeta^{\bullet,k}(t_n)\right)}{n} \xrightarrow{\text{a.s.}} \frac{\log(k!)}{k!}\int_0^1\int_0^x \frac{y^k e^{-y}}{(1-y)}dy dx. \tag{5.3.3}$$

## 5.4 Recursively exponential polynomials

In Proposition 5.3.7 we proved that $\lim_{n\to\infty}\frac{\log(\zeta^{\bullet,k})}{n}$ can be calculated by evaluating an integral. In this section we develop an algebraic toolbox necessary to calculate the this integral (the right-hand side of Equation 5.3.3) using a novel family of polynomials called recursively exponential polynomials. We end this section by calculating $\mathcal{L}_1$.

For succinctness, given a polynomial $p(x)$, we define

$$\frac{d^0}{dx^0}p(x) := p(x).$$

Given such a polynomial $p(x)$, consider $\int e^{-x}p(x)dx$. Integrating by parts we see that,

$$\int e^{-x}p(x)dx = -e^{-x}p(x) + \int e^{-x}\frac{d}{dx}p(x)dx$$

and by iterating this argument we see that

$$\int e^{-x}p(x)dx = -e^{-x}\left(\sum_{j\geq 0}\frac{d^j}{dx^j}p(x)\right). \tag{5.4.1}$$

**Definition 5.4.1.** (Recursively exponential polynomials). Define a family $\{p_k(x)\}_{k\in\mathbb{N}}$ of polynomials in one variable by $p_1(x) := 1$, then recursively by

$$p_{k+1}(x) := 1 + xp_k(x) + \sum_{j\geq 0}\frac{d^j}{dx^j}p_k(x).$$

**Example 5.4.2.** The first few recursively exponential polynomials are,

$$\begin{aligned}
p_1(x) &= 1\\
p_2(x) &= x + 2\\
p_3(x) &= x^2 + 3x + 4\\
p_4(x) &= x^3 + 4x^2 + 9x + 10.
\end{aligned}$$

**Proposition 5.4.3.** For all $k \in \mathbb{N}$ we have the following:

$$\text{(i) } p_k(0) = \sum_{i=0}^{k-1} i!$$

$$\text{(ii) } k!\sum_{i=0}^{k}\frac{1}{i!} = \sum_{j\geq 0}\frac{d^j}{dx^j}p_k(x)\bigg|_{x=1} + 1.$$

Before we prove Proposition 5.4.3 it is necessary to introduce another new family of polynomials, $q_k(x) := p_{k+1}(x) - p_k(x)$, that we call *difference polynomials*. We immediately remark that

$$\begin{aligned}
q_k(x) &= 1 + xp_k(x) + \sum_{j\geq 1}\frac{d^j}{dx^j}p_k(x)\\
&= 1 + x(p_{k-1}(x) + q_{k-1}(x)) + \sum_{j\geq 1}\frac{d^j}{dx^j}p_{k-1}(x) + q_{k-1}(x)\\
&= xq_k(x) + \sum_{j\geq 1}\frac{d^j}{dx^j}q_k(x) + 1 + xp_k(x) + \sum_{j\geq 1}\frac{d^j}{dx^j}p_k(x)\\
&= xq_k(x) + \sum_{j\geq 0}\frac{d^j}{dx^j}q_k(x)
\end{aligned} \tag{5.4.2}$$

for all $k \in \mathbb{N}$. For each differernce polnomial we write $q_k(x) = \sum_{i=0}^{k} \lambda_{k,i} x^i$ for some constants $\lambda_{k,i} \in \mathbb{N}$.

**Lemma 5.4.4.** Fix $k \in \mathbb{N}$. Then,

$$\lambda_{k,i} = \frac{k!}{i!}$$

for all $i \leq k$ and $\lambda_{k,i} = 0$ otherwise.

*Proof.* We use induction on $k$. For the base case we remark that $q_1(x) = x + 1$.

We begin with the case $i \in \mathbb{N}$ and then consider the case $i = 0$. So suppose $q_k(x) = \sum_{i=0}^{k} \frac{k!}{i!} x^i$ for $k = i, i+1, \ldots, K$. By Equation 5.4.2 and then the inductive hypothesis,

$$\begin{aligned}
\lambda_{K+1,i} &= \lambda_{K,i-1} + \sum_{j=0}^{K-i} \frac{(i+j)!}{i!} \lambda_{K,i+j} \\
&= \lambda_{K,i-1} + \sum_{j=0}^{K-i} \frac{(i+j)!}{i!} \lambda_{K,i+j} \\
&= \frac{K!}{(i-1)!} + \sum_{j=0}^{K-i} \frac{(i+j)!K!}{i!(i+j)!} \\
&= \frac{(K+1)!}{i!}
\end{aligned}$$

for each $i \in [k]$.

Now consider the case $i = 0$. By the inductive hypothesis,

$$\begin{aligned}
\lambda_{K+1,0} &= \sum_{j=0}^{K} j! \lambda_{K,j} \\
&= \sum_{j=0}^{K} \frac{j!K!}{j!} \\
&= K + 1!
\end{aligned}$$

and the result follows. □

*Proof.* (of Proposition 5.4.3)

(i) We prove this by induction on $k$. For the base case we note that $p_1(x) = 1$. Suppose, for the inductive hypothesis, $p_k(0) = \sum_{i=0}^{k-1} i!$ for $k = 1, 2, \ldots, K$. We remark that $p_{K+1}(0) - p_K(0) = q_K(0)$. By the inductive hypothesis, $p_{K+1}(0) = q_K(0) + \sum_{i=0}^{K-1} i!$. It follows immediately from Lemma 5.4.4 that $q_K(0) = K!$, hence

$$p_{K+1}(0) = K! + \sum_{i=0}^{K-1} i!$$

as required.

(ii) We remark that

$$\begin{aligned}
k!\sum_{i=0}^{k}\frac{1}{i!} &= \sum_{j\geq 0}\frac{d^j}{dx^j}p_k(x)\Big|_{x=1}+1 \\
&= p_{k+1}(1)-p_k(1) \\
&= q_k(1) \\
&= \sum_{i=0}^{k}\frac{k!}{i!}
\end{aligned}$$

by Lemma 5.4.4.

□

**Definition 5.4.5** (exponential integral)**.** [Vau08]. We define the *exponential integral*

$$\operatorname{Ei}(x) := \int_{-x}^{\infty}\frac{e^{-t}}{t}dt.$$

which should be interpreted in the sense of Cauchy principal value.

By integrating by parts we get,

$$\int \operatorname{Ei}(x)dx = x\operatorname{Ei}(x) - e^x + C. \tag{5.4.3}$$

Therefore,

$$\int \operatorname{Ei}(1-x)dx = e^{1-x} - (1-x)\operatorname{Ei}(1-x) + C \tag{5.4.4}$$

by substitution.

**Lemma 5.4.6.** Suppose $k \in \mathbb{N}$. Then,

$$G_k(x) = e^{-x}(p_k(x)) - \frac{\operatorname{Ei}(1-x)}{e} - p_k(0) + \frac{\operatorname{Ei}(1)}{e}.$$

We prove Lemma 5.4.6 by induction on $k$. Before we prove Lemma 5.4.6 we prove the base case $k = 1$.

**Lemma 5.4.7.**

$$\int_0^x \frac{ye^{-y}}{(1-y)}dy = e^{-x} - \frac{\operatorname{Ei}(1-x)}{e} - p_1(0) + \frac{\operatorname{Ei}(1)}{e}.$$

*Proof.* First note that

$$\int \frac{xe^{-x}}{(1-x)}dx = -\int \frac{xe^{-x}}{(x-1)}dx,$$

now we use integration by parts to see that

$$\int \frac{xe^{-x}}{(x-1)}dx = \frac{x\,\mathrm{Ei}(1-x)}{e} - \frac{1}{e}\int \mathrm{Ei}(1-x)dx.$$

By Equation 5.4.4 we get

$$\begin{aligned}\int_0^x \frac{ye^{-y}}{(y-1)}dy &= \left[\frac{\mathrm{Ei}(1-y)}{e} - e^{-y}\right]_0^x \\ &= \frac{\mathrm{Ei}(1-x)}{e} - e^{-x} - \frac{\mathrm{Ei}(1-x)}{e} + 1\end{aligned}$$

and the result follows. □

*Proof.* (of Lemma 5.4.6) Suppose, for the inductive hypothesis, that

$$\int \frac{x^n e^{-x}}{(1-x)}dx = e^{-x}(p_n(x)) - \frac{\mathrm{Ei}(1-x)}{e}$$

for $n = 1, 2, \ldots, k-1 > 0$. Integrating by parts and applying Equations 5.4.1 and 5.4.4 we see that

$$\begin{aligned}\int \frac{x^k e^{-x}}{(x-1)}dx &= \int x\left(\frac{x^{k-1}e^{-x}}{(x-1)}\right)dx \\ &= x\left(\frac{\mathrm{Ei}(1-x)}{e} - e^{-x}p_{k-1}(x)\right) - \int \frac{\mathrm{Ei}(1-x)}{e}dx - e^{-x}p_{k-1}(x)dx \\ &= \frac{\mathrm{Ei}(1-x)}{e} - e^{-x}\left(xp_{k-1}(x) + \sum_{j\geq 0}\frac{d^j}{dx^j}p_{k-1}(x) + 1\right) \\ &= \frac{\mathrm{Ei}(1-x)}{e} - e^{-x}p_k(x)\end{aligned}$$

where the last equality follows from Definition 5.4.1. Furthermore,

$$\begin{aligned}\int_0^x \frac{y^n e^{-y}}{(1-y)}dy &= \left[e^{-y}(p_n(y)) - \frac{\mathrm{Ei}(1-y)}{e}\right]_0^x \\ &= \frac{\mathrm{Ei}(1-x)}{e} - e^{-x}p_k(x) - \frac{\mathrm{Ei}(1)}{e} + p_k(0).\end{aligned}$$

□

**Corollary 5.4.8.** For all $k \in N$,

$$H_k(x) = -e^{-x}\left(\sum_{i\geq 0}\frac{d^i}{dx^i}p_k(x) - 1\right) + \frac{(1-x)}{e}\left(\mathrm{Ei}(1-x) - \mathrm{Ei}(1)\right) - xp_k(0) + p_{k+1}(0).$$

*Proof.* For all $k \in \mathbb{N}$,

$$\begin{aligned} H_k(x) &= \int_0^x G_k(y)dy \\ &= \int_0^x \left( e^{-y}(p_k(y)) - \frac{\operatorname{Ei}(1-y)}{e} - p_k(0) + \frac{\operatorname{Ei}(1)}{e} \right) dy. \end{aligned}$$

By Equation 5.4.1,

$$\int e^{-x} p_k(x) dx = -e^{-x} \left( \sum_{j \geq 0} \frac{d^j}{dx^j} p_k(x) \right) \tag{5.4.5}$$

and, by Equation 5.4.4,

$$\int \frac{\operatorname{Ei}(1-x)}{e} dx = \frac{1}{e} \left( e^{1-x} - (1-x) \operatorname{Ei}(1-x) \right). \tag{5.4.6}$$

The result follows. □

We finally have enough algebraic machinery to prove Theorem 5.4.9.

**Theorem 5.4.9.** Suppose that $\{t_n\}_{n \in \mathbb{N}}$ is a random recursive tree. Then,

$$\zeta^{\bullet}(t_n)^{\frac{1}{n}} \xrightarrow{\text{a.s.}} \prod_{k \geq 2} (k!)^{\gamma_k}$$

where $\gamma_k = 1 - \frac{1}{e} \sum_{j=0}^{k} \frac{1}{j!}$ for $k \geq 2$.

*Proof.* By Theorem 5.3.6,

$$\frac{\log \zeta^{\bullet, k}(t_n)}{n} \xrightarrow{\text{a.s.}} \frac{\log(k!)}{k!} H_k(1)$$

for any $k \in \mathbb{N}$. By Corollary 5.4.8,

$$\begin{aligned} H_k(1) &= \frac{-1}{e} \left[ \sum_{i \geq 0} \frac{d^i}{dx^i} p_k(x) \right]_{x=1} - \frac{1}{e} + \left( \frac{\operatorname{Ei}(1)}{e} - p_k(0) \right) - \frac{\operatorname{Ei}(1)}{e} + p_{k+1}(0) \\ &= p_{k+1}(0) - p_k(0) - \frac{1}{e} \left( k! \sum_{i=0}^{k-1} \frac{1}{i!} + 1 \right) \end{aligned}$$

by Proposition 5.4.8 Part (iii). By Proposition 5.4.8 Part (i), $p_k(0) = \sum_{i=0}^{k-1} i!$, for all $k \in \mathbb{N}$. Hence,

$$\begin{aligned} H_k(1) &= \sum_{i=0}^{k} i! - \sum_{i=0}^{k-1} i! - \frac{1}{e}\left(k!\sum_{i=0}^{k-1}\frac{1}{i!} + 1\right) \\ &= k! - \frac{1}{e}\left(k!\sum_{i=0}^{k-1}\frac{1}{i!} + 1\right) \\ &= k!\left(1 - \frac{1}{e}\sum_{i=0}^{k}\frac{1}{i!}\right) \end{aligned}$$

Suppose $\{t_n\}_{n\in\mathbb{N}} \in \mathcal{T}_{\tilde{w}}$. Then, by Theorem 5.3.6,

$$\begin{aligned} \frac{\log \zeta^{\bullet,k}(t_n)}{n} &\xrightarrow{\text{a.s.}} H_k(1)\frac{\log(k!)}{k!} \\ &= \log(k!)\left(1 - \frac{1}{e}\sum_{i=0}^{k}\frac{1}{i!}\right) \end{aligned}$$

for any $k \in \mathbb{N}$. The result follows since, $\mathrm{Aut}^\bullet(t) = \prod_k \mathrm{Aut}^{\bullet,k}(t)$. □

We used Mathematica [Wol16] to calculate,

$$\sum_{k=2}^{1000} \log(k!)\left(1 - \frac{1}{e}\sum_{i=0}^{k}\frac{1}{i!}\right) \approx 0.10480730877\ldots \tag{5.4.7}$$

Hence, by the Continuous Mapping Theorem (Theorem 3.2.3),

$$\lim_{n\to\infty}\sum_{k=2}^{1000}\zeta^{\bullet,k}(t_n)^{\frac{1}{n}} \approx 1.11049660678 =: \mathcal{L}_1$$

almost surely. In this way it is possible to get a numerical estimate for the limiting order for the subgroup of the automorphism group of a random recursive tree coming from $k$-stars. We remark that since $\mathcal{L}_1 < \mathcal{V}$ we have not (yet) proved Theorem 3.

## 5.5 Proof of Theorem 3

In this section we will calculate a lower bound, $\mathcal{L}_2$, for $\prod_{s\in\mathcal{R}\setminus\bullet}\zeta^s(t_n)^{\frac{1}{n}}$. At the end of this section we prove that $\mathcal{L}_1\mathcal{L}_2 > \mathcal{V}$ (Viswanath's constant) thus proving Theorem 3.

Following Pitman [Pit95] we put a measure of $\Sigma_n$ as follows. Initially $\sigma_1(1) = 1$. At time $n > 1$,

$$\mathbb{P}\left(\sigma_n(n) = j\right) = \frac{1}{n}$$

for any $j \in [n]$. This corresponds, under $\widetilde{T}$, to the symmetric group equipt with the uniform measure which we denote $S_n^{\mu}$.

Fix $\sigma \in S_n$ and suppose $C = (c_1, c_2, \dots, c_r)$ is a $r$-cycle of $\sigma$. Let $X_C = \{x \in \mathbb{N} : x = c_i \text{ for some } i \in [r]\}$ . Order $X_C = \{x_1, x_2, \dots, x_r\}$ so that $x_1 < x_2 < \cdots < x_r$. Define a map $\iota : X_C \to [r]$ by $\iota(x_i) := i$ for each $i \in [r]$.

Define a function $\widehat{T}|_C : S_n^{\mu} \to \mathcal{T}_{w,r}$ by

$$\widehat{T}|_C(\sigma) := (v_2, \dots, v_r)$$

where $v_i = \iota(\sigma_{x_i}(x_i))$ for $i = 2, 3, \dots, r$ and $\widetilde{T}\left(\{\sigma_i\}_{i=1}^n\right) = \sigma$. Suppose $\sigma \in S_n^{\mu}$. Then,

$$\mathbb{P}\left(\sigma_{x_j}(x_j) = x_i | \sigma_{x_j}(x_j) \in \{x_1, x_2, \dots, x_j\}\right) = \frac{1}{j}$$

for any pair $j \in [r]$ and $i \in [j]$. In particular $\widehat{T}|_C : S_n^{\mu} \to \mathcal{T}_{\widetilde{w},r}$.

**Example 5.5.1.** The permutation $\sigma = (3, 4, 1)(5, 6, 2)$ corresponds, under the map $\widehat{T}$, to the attachment tree $\{t_i\}_{i=1}^7$ shown in Figure 5.1. Label the cycles of $\sigma$ as follows,

$$C_1 = (3, 4, 1) \text{ and } C_2 = (5, 6, 2).$$

Then,

$$\widehat{T}|_{C_1}(\sigma) = (1, 1) \text{ and } \widehat{T}|_{C_2}(\sigma) = (1, 1)$$

which is the attachment tree shown in Figure 5.2.

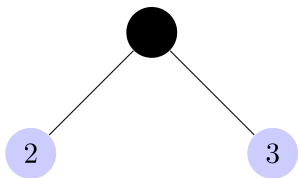


Figure 5.2: A representation of random recursive tree with vertex sequence $(1, 1)$.

Recall from Section 3.1 that $\phi_n : \mathcal{T}_{w,n} \to \mathcal{R}_n$ is the map that simply "forgets" the random recursive structure of a $n$-sapling. We have shown that:

**Lemma 5.5.2.** Fix $r \in \mathbb{N}$ and a rooted tree $s \in \mathcal{R}_r$ and suppose $C$ is a $r$-cycle of $\sigma \in S_n$. Under the uniform distribution on $S_n$,

$$\mathbb{P}\left(\phi_n\left(\widehat{T}|_C(\sigma)\right) = s\right) = \frac{K(s)}{(r-1)!}$$

Let $\sigma \in S_n$. We write $C^s(\sigma)$ for the number of cycles $C$ of $\sigma$ such that

$$\phi_{|C|}\left(\widehat{T}|_C(\sigma)\right) = s.$$

Suppose $\sigma$ has cycle type $\mathbf{a}$. By Lemma 5.5.2,

$$\mathbb{P}\left(C^s(\sigma) = a_r\right) = \left(\frac{K(s)}{(r-1)!}\right)^{a_r}.$$

Define,

$$A_i^k := \sum_{s \in \mathcal{R}_i} \left(\frac{K(s)}{(i-1)!}\right)^k$$

**Example 5.5.3.** We calculate $A_i^k$ for $i \leq 4$. For all $k \geq 1$,

$$\begin{aligned} A_1^k &= 1 \\ A_2^k &= 1 \\ A_3^k &= \left(\frac{1}{2}\right)^{k-1} \\ A_4^k &= \left(\frac{1}{2}\right)^k + 3\left(\frac{1}{6}\right)^k. \end{aligned}$$

Let $\sigma \in S$. We write $C_i^s(\sigma)$ for the number of $i$-cycles, $C$, of $\sigma_n$ such that

$$\phi_{|C|}\left(\widehat{T}|_C(\sigma)\right) = s.$$

Define $S_n^i \subseteq S_n$ to be the set of permutations $\sigma \in S_n$ such that

$$C_i^s(\sigma) = a_i.$$

In other words under the bijection $\widehat{T}$ the subset $S_n^i$ corresponds to all random recursive trees consisting of a root vertex attached to a rooted forest such that every rooted tree in that rooted forest on $i$ vertices is isomorphic. Define $\widehat{X}_n^{i,k}$ to be the number of permutations $\sigma \in S_n^i$ with precisely $k$ cycles of length $i$. Note that,

$$\widehat{X}_n^{i,k} = X_n^{i,k} A_i^k.$$

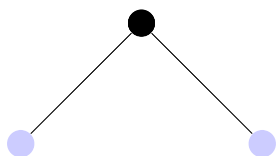

Figure 5.3: A rooted tree $s \in \mathcal{R}$. We remark that $s = \phi_3((1,1))$, depicted in Figure 5.2. Further, $C^s(\sigma) = 2$ where $\sigma = \widetilde{T}^{-1}(\{\sigma_i\}_{i=1}^6)$ is as in Example 5.1.

Given a permutation $\sigma$ such that $\widehat{T}(\sigma) = \{t_i\}_{i=1}^{n+1}$, we have

$$\log(\Lambda(t_{n+1})) = \sum_{s\in\mathcal{R}} \log(C^s(\sigma)!), \tag{5.5.1}$$

Suppose $\sigma \in S_n$. Let $\tilde{j} : S_n \to \mathbb{R}$ be the function defined by,

$$\tilde{j}(\sigma) := \sum_{s\in\mathcal{R}} \log(C^s(\sigma)!),$$

Recall from the introduction to this Chapter that calculating such a function $\tilde{j}$ is integral to achieving our long term aim of proving Theorem 3.

**Lemma 5.5.4.**

$$\log(\mathcal{W}) \geq \sum_{n\geq 0}\sum_{i\geq 1}\sum_{k=2}^{n} \frac{\widehat{X}_n^{i,k}\log(k!)}{(n+2)!}. \tag{5.5.2}$$

*Proof.* By Equation 5.5.1,

$$\begin{aligned}
\sum_{t\in\mathcal{R}_{n+1}} K(t)\log(\Lambda(t)) &= \sum_{\sigma\in S_n}\sum_{r\in\mathcal{R}} \log\left(C^r(\sigma)!\right) \\
&= \sum_{\sigma\in S_n}\sum_{i=1}^{n}\sum_{s\in\mathcal{R}_i} \log\left(C_i^s(\sigma)!\right) \\
&= \sum_{i=1}^{n}\sum_{\sigma\in S_n}\sum_{s\in\mathcal{R}_i} \log\left(C_i^s(\sigma)!\right) \\
&\geq \sum_{i=1}^{n}\sum_{\sigma\in S_n^i}\sum_{s\in\mathcal{R}_i} \log\left(C_i^s(\sigma)!\right) \\
&= \sum_{i=1}^{n}\sum_{\sigma\in S_n^i} \log(a_i!) \\
&= \sum_{i=1}^{n}\sum_{k=2}^{n} \widehat{X}_n^{i,k}\log(k!).
\end{aligned}$$

The result follows. □

In Lemma 5.5.5 we simplify the righthand side of inequality 5.5.2 in order that we might more easily approximate it using a computer. In [Sta86] Stanley proves the number $X_n^{i,k}$ of permutations $\sigma \in S_n$ such that $a_i = k$ is

$$X_n^{i,k} := n! \sum_{j=0}^{\lfloor \frac{n}{i} - k \rfloor} \frac{(-1)^j}{i^j j!}.$$

**Lemma 5.5.5.** Fix $i, k \in \mathbb{N}$. Then,

$$\sum_{n \geq ik} \frac{X_n^{i,k}}{(n+2)!} = \sum_{m \geq 0} \frac{i \sum_{j=0}^{m} \frac{(-1)^j}{i^j j!}}{(i(k+m)+1)(i(k+m+1)+1)}$$

*Proof.* For every $i \geq 1$,

$$\begin{aligned}\sum_{n \geq ik} \frac{1}{(n+1)(n+2)} \sum_{j=0}^{\lfloor \frac{n}{i} - k \rfloor} \frac{(-1)^j}{i^j j!} &= \frac{\sum_{j=0}^{0} \frac{(-1)^j}{i^j j!}}{(ik+1)(ik+2)} + \frac{\sum_{j=0}^{0} \frac{(-1)^j}{i^j j!}}{(ik+2)(ik+3)} + \cdots + \frac{\sum_{j=0}^{m} \frac{(-1)^j}{i^j j!}}{(ik+i)(ik+i+1)} \\ &+ \frac{\sum_{j=0}^{1} \frac{(-1)^j}{i^j j!}}{(i(k+1)+1)(i(k+1)+2)} + \cdots + \frac{\sum_{j=0}^{1} \frac{(-1)^j}{i^j j!}}{(i(k+1)+i)(i(k+1)+i+1)} \\ &+ \frac{\sum_{j=0}^{2} \frac{(-1)^j}{i^j j!}}{(i(k+2)+1)(i(k+2)+2)} + \ldots\end{aligned}$$

for any $k \geq 2$. Since $\sum_{n=0}^{N} \frac{1}{(n+1)(n+2)} = \frac{N+1}{N+2}$ we have,

$$\sum_{A}^{B} \frac{1}{(n+1)(n+2)} = \frac{B-A+1}{(A+1)(B+2)}$$

for any $A, B \in \mathbb{N}$. In particular,

$$\sum_{n \geq ik} \frac{1}{(n+1)(n+2)} \sum_{j=0}^{\lfloor \frac{n}{i} - k \rfloor} \frac{(-1)^j}{i^j j!} = \sum_{m \geq 0} \frac{i \sum_{j=0}^{m} \frac{(-1)^j}{i^j j!}}{(i(k+m)+1)(i(k+m+1)+1)}.$$

□

**Corollary 5.5.6.** For $I \in \mathbb{N}, M \in \mathbb{N}$ and $2 \leq K \in \mathbb{N}$,

$$\log(\mathcal{W}) \geq \sum_{i=1}^{I} \sum_{k=2}^{K} \sum_{m=0}^{M} \frac{i A_i^k \sum_{j=0}^{m} \frac{(-1)^j}{i^j j!}}{(i(k+m)+1)(i(k+m+1)+1)}$$

*Proof.* Let $M' = I(K+M+1)-1$. By the proof of Lemma 5.5.5,

$$\begin{aligned}\sum_{i=1}^{I} \sum_{k=2}^{K} \sum_{m=0}^{M} \frac{i A_i^k \sum_{j=0}^{m} \frac{(-1)^j}{i^j j!}}{(i(k+m)+1)(i(k+m+1)+1)} &= \sum_{i=1}^{I} \sum_{k=2}^{K} \sum_{n=0}^{i(k+m+1)-1} \frac{\widehat{X}_n^{i,k} \log(k!)}{(n+2)!} \\ &= \sum_{n=0}^{M'} \sum_{i=1}^{I} \sum_{k=2}^{K} \frac{\widehat{X}_n^{i,k} \log(k!)}{(n+2)!} \\ &\leq \sum_{n \geq 0} \sum_{i \geq 1} \sum_{k \geq 2} \frac{\widehat{X}_n^{i,k} \log(k!)}{(n+2)!} \\ &\leq \log(\mathcal{W})\end{aligned}$$

□

We used Mathematica [Wol16] to calculate,

$$\sum_{i=2}^{4}\sum_{k=2}^{6}\sum_{m=0}^{400}\frac{iA_i^k\sum_{j=0}^{m}\frac{(-1)^j}{i^j j!}}{(i(k+m)+1)(i(k+m+1)+1)} \approx 0.0195308138\ldots$$
$$= \log(\mathcal{L}_2)$$

*Proof.* (of Theorem 3.) Note that

$$\mathcal{L}_1\mathcal{L}_2 = 1.132398696\ldots$$

By Corollary 5.5.6 $\mathcal{V} < \mathcal{L}_1\mathcal{L}_2 \leq \mathcal{W}$ and the result follows. $\square$

# Chapter 6

# Expectation

Whilst the expected order of the automorphism group of certain families of rooted trees such as labelled rooted trees are known [Yu12], there are no analogous formulae for other important families of trees. In this chapter we calculate the expected order $b_k$ of the automorphism group of a $k$-sapling $\{t_i\}_{i=1}^{k} \in \mathcal{T}_{\widetilde{w}}$ for all $k \in \mathbb{N}$. Our primary tool for proof in this chapter is generating functions.

The main result of this chapter is Theorem 4. We say that Theorem 4 is a *Fibonacci-type* theorem because of the obvious similarity between Equation 1.4.2 and Binet's formula for Fibonacci numbers. Theorem 4 is particularly astonishing because although the expected order of the automorphism group of a rooted labelled tree is known, it does not have a closed form expression.

The key to proving Theorem 4 is Equation 3.2.3 which was derived by Hoffman in [Hof03] and gives an easily manipulated expression for the expected order of the automorphism group of a random recursive tree. Hoffman's derivation of Equation 3.2.3 relies on his observation that the set of rooted trees with a finite number of vertices satisfies the definition of a partially ordered set with certain algebraic and combinatorial properties called a *sequentially differential poset* which will be defined in Section 6.2.

Sequentially differential posets were first introduced by Stanley in [Sta90] as a generalisation of another family of posets called *differential posets* which share many of the same combinatorial properties. Applications of differential posets are numerous and include the study of Hopf Algebras [BLL12, Hof03], Markov processes [Ful09] and Young-Fibonacci lattices. The combinatorics of differential posets first outlined in [Sta88a] can be expressed elegantly in terms of generating functions whereas the combinatorics of sequentially differential posets cannot and thus become much more complicated. In Section 6.2 we show that the combinatorics of a family of sequentially differential posets called *sequentially increasing posets* (to which the set of rooted trees with finitely many vertices belongs) do admit a generating function description.

## 6.1 A generating function approach

An *involution* $\sigma \in S_k$ does not contain any permutation cycles of length $> 2$ so that it consists exclusively of fixed points and transpositions with disjoint supports. The set of all involutions of $[k]$ is denoted $\mathrm{Inv}(k)$. For $\sigma \in S_k$, call $i \in [k]$ a *weak excedance* of $\sigma$ if $\sigma(i) \geq i$. Let $\mathrm{Wex}(\sigma)$ be the set of weak excedances of $\sigma$. For $\sigma \in S_k$ and $i \in [k]$, let $\eta(\sigma, i)$ be the number of integers $j \in [k]$ such that $j < i$ and $\sigma(j) < \sigma(i)$. With the definitions above Hoffman [Hof03] proved

$$\sum_{t \in \mathcal{T}_{\widetilde{w},k+1}} \zeta(t) = \sum_{\sigma \in \mathrm{Inv}(k)} \prod_{i \in \mathrm{Wex}(\sigma)} (\eta(\sigma, i) + 1). \tag{6.1.1}$$

In this section we will define a generating function that describes Equation 6.1.1, hence we define

$$e_k = \sum_{\sigma \in \mathrm{Inv}(k)} \prod_{i \in \mathrm{Wex}(\sigma)} (\eta(\sigma, i) + 1).$$

In Lemma 6.1 we will prove that the sequence $\{e_k\}_{k \geq 0}$ can be defined in terms of a simple recursion redolent of Fibonacci's sequence.

**Lemma 6.1.** *The sequence $\{e_k\}_{k \geq 0}$ is given by the recursion*

$$e_k = ke_{k-1} + \frac{k(k-1)e_{k-2}}{2},$$

*for $k \geq 2$ and has initial values $e_0 = e_1 = 1$.*

*Proof.* Fix $k \in \mathbb{N}$ and suppose that $\sigma \in \mathrm{Inv}(k)$. Either $k$ is a fixed point or $k$ is part of a transposition.

**Case 1.** Suppose that $k$ is a fixed point of $\sigma$, then we can split $\sigma$ into a map that sends $k$ to $k$ and a permutation $\tau \in \Sigma_{k-1}$ which simply forgets $k$ so that $\tau(i) = \sigma(i)$ for all $i \in [k-1]$. Since $\sigma(k) = k$, $k$ is a weak excedance of $\sigma$. For all $j \in [k-1]$ we have $j < k$ and $\sigma(j) < k = \sigma(k)$ hence $\eta(\sigma, k) = k - 1$.

Now we will consider $\eta(\sigma, i)$ for each $i < k$. It is clear that $i \in \mathrm{Wex}(\sigma)$ if and only if $i \in \mathrm{Wex}(\tau)$ for all $i \in [k-1]$. Since $\tau(i) = \sigma(i)$ for all $i \in [k-1]$, $\eta(\tau, i) = \eta(\sigma, i)$ for all $i \in [k-1]$, hence

$$\begin{aligned}\prod_{i \in Wex(\sigma)} (\eta(\sigma, i) + 1) &= (\eta(\sigma, k) + 1) \left( \prod_{i \in Wex(\tau)} (\eta(\tau, i) + 1) \right) \\ &= k \left( \prod_{i \in Wex(\tau)} (\eta(\tau, i) + 1) \right).\end{aligned}$$

We remark that $\sigma \in \mathrm{Inv}(k)$ if and only if $\tau \in \mathrm{Inv}(k-1)$, hence in this case $e_k = ke_{k-1}$.

**Case 2.** Now suppose that $k$ is part of a transposition, $(j, k)$, for some $j \in [k-1]$. Define a new set $[k-1]_j := \{1, 2, \ldots, j-1, j+1, \ldots, k-1\}$. Again we split $\sigma$ into a map that sends $j$ to $k$ (and $k$ to $j$) and a map $\tilde{\rho} : [k-1]_j \longrightarrow [k-1]_j$ given by $\tilde{\rho}(i) = \sigma(i)$ for all $i \in [k-1]_j$. In order to think of $\tilde{\rho}$ as a permutation we define a function $f : [k-1]_j \to [k-2]$ as follows:

$$f(i) = \begin{cases} i & \text{if } i < j \\ i-1 & \text{if } i > j. \end{cases}$$

This function is clearly a bijection. Define a new permutation $\rho \in \Sigma_{k-2}$ by $\rho(i) := f(\tilde{\rho}(f^{-1}(i))$.

Note that $j \in \mathrm{Wex}(\sigma)$ and $k \notin \mathrm{Wex}(\sigma)$ and that $\eta(\sigma, j) = j-1$. Again notice that $\eta(\sigma, f^{-1}(i)) = \eta(\rho, i)$ for all $i \in [k-2]$. Since $i \in \mathrm{Wex}(\rho)$ if and only if $f^{-1}(i) \in \mathrm{Wex}(\sigma)$ for all $i \in [k-2]$,

$$\prod_{i \in \mathrm{Wex}(\sigma)} (\eta(\sigma, i) + 1) = j \left( \prod_{i \in \mathrm{Wex}(\rho)} (\eta(\rho, i) + 1) \right).$$

We remark that $\sigma \in \mathrm{Inv}(k)$ if and only if $\rho \in \mathrm{Inv}(k-2)$ hence in this case $e_k = je_{k-2}$ for a fixed $j \in [k-1]$.

We combine Case 1 and Case 2 to see that

$$\begin{aligned} e_k &= ke_{k-1} + \sum_{j=1}^{k-1} je_{k-2} && (6.1.2) \\ &= ke_{k-1} + \frac{k(k-1)}{2} e_{k-2}. && (6.1.3) \end{aligned}$$

By the definition of $e_k$ it is easy to see that $e_0 = e_1 = 1$. $\square$

Equation 6.1.3 naturally lends itself to a generating function argument so we define the exponential generating function $F(x) = \sum_{k \geq 0} \frac{e_k x^k}{k!}$. Following a standard generating function type argument:

$$\begin{aligned} F(x) &= 1 + x + \sum_{k \geq 2} \frac{e_k x^k}{k!} && (6.1.4) \\ &= 1 + x + \sum_{k \geq 2} \left( ke_{k-1} + \frac{k(k-1)}{2} e_{k-2} \right) \frac{x^k}{k!} && (6.1.5) \\ &= 1 + x + x(F(x) - 1) + \frac{x^2}{2} F(x). && (6.1.6) \end{aligned}$$

Rearranging Equation 6.1.6 for $F(x)$ we see that the exponential generating function for the recursion given in Equation 6.1.3 is

$$F(x) = \frac{2}{(2 - 2x - x^2)}. \tag{6.1.7}$$

**Remark 6.1.1.** The Fibonacci sequence, $\{f_i\}_{i\geq 0}$, is given by $f_0 = f_1 = 1$ and subsequently $f_k = f_{k-1} + f_{k-2}$ for $k > 2$. Define the generating function $G(x) := \sum_{k\geq 0} f_k x^k$ and an elementary argument gives

$$G(x) = \frac{1}{(1 - x - x^2)}.$$

There is an obvious similarity between $F(x)$ and $G(x)$. A closed form for the $k^{th}$ term of Fibonacci's sequence, $f_k$, can be calculated by a standard generating function argument and yields the following result:

$$f_k = \frac{1}{\sqrt{5}} \left( \left( \frac{1+\sqrt{5}}{2} \right)^k - \left( \frac{1-\sqrt{5}}{2} \right)^k \right), \tag{6.1.8}$$

known as Binet's formula.

In the proof of Theorem 4 we will adapt this standard argument to give a closed form for the expected order of the automorphism group of a random recursive tree.

*Proof of Theorem 4.* The function $f(x) = x^2 + 2x - 2$ has solutions $\phi_1 = -1 + \sqrt{3}$ and $\phi_2 = -1 - \sqrt{3}$, hence

$$F(x) = \frac{-2}{(\phi_1 - x)(\phi_2 - x)}$$

By partial fractions then a Taylor expansion:

$$\begin{aligned}
F(x) &= \frac{2}{\phi_1 - \phi_2} \left( \frac{1}{\phi_1 - x} - \frac{1}{\phi_2 - x} \right) && (6.1.9)\\
&= \frac{2}{\phi_1 - \phi_2} \left( \frac{\phi_1^{-1}}{1 - \frac{x}{\phi_1}} - \frac{\phi_2^{-1}}{1 - \frac{x}{\phi_2}} \right) && (6.1.10)\\
&= \frac{2\phi_1^{-1}}{\phi_1 - \phi_2} \left( \frac{1}{1 - \frac{x}{\phi_1}} \right) - \frac{2\phi_2^{-1}}{\phi_1 - \phi_2} \left( \frac{1}{1 - \frac{x}{\phi_2}} \right) && (6.1.11)\\
&= \frac{2\phi_1^{-1}}{\phi_1 - \phi_2} \left( \sum_{k\geq 0} \frac{x^k}{\phi_1^k} \right) - \frac{2\phi_2^{-1}}{\phi_1 - \phi_2} \left( \sum_{k\geq 0} \frac{x^k}{\phi_2^k} \right) && (6.1.12)\\
&= 2 \sum_{k\geq 0} \left( \frac{\phi_1^{-k-1} - \phi_2^{-k-1}}{\phi_1 - \phi_2} \right) x^k. && (6.1.13)
\end{aligned}$$

We define $b_k = \frac{e_k}{k!}$. By Equation 6.1.13 $b_k = 2\left(\frac{\phi_1^{-k-1} - \phi_2^{-k-1}}{\phi_1 - \phi_2}\right)$. Since $\phi_1^{-1} = (1+\sqrt{3})/2$ and $\phi_2^{-1} = (1-\sqrt{3})/2$,

$$b_k = \frac{1}{\sqrt{3}}\left(\left(\frac{1+\sqrt{3}}{2}\right)^{k+1} - \left(\frac{1-\sqrt{3}}{2}\right)^{k+1}\right).$$

Since $\left|\mathcal{T}_{\widetilde{w},k+1}\right| = k!$ the result follows. $\square$

## 6.2 Differential posets

In Section 6.1 we remarked that the set of rooted trees on a finite number of vertices satisfies the axioms of a particular kind of partially ordered set called a differential poset. In this section we give the formal definition of a differential poset and in Proposition 6.2.2 we use the machinery of differential posets to show that Theorem 4 can be generalised. We begin, as ever, with a few definitions.

A poset $P$ is called *locally finite* if for all $x, y \in P$ the interval $[x, y]$ consists of finitely many elements. Suppose that $P$ is a poset with partial ordering $\preceq$, then we write $x \prec y$ for any $x, y \in P$ such that $x \preceq y$ and $x \neq y$. We say that $y$ *covers* $x$ if $x \prec y$ and there does not exist an element $z \in P$ such that $x \prec z \prec y$. We also say that $x$ is *covered* by $y$. A poset $P$ is said to be *graded* if $P$ is equipped with a function $\rho : P \to \mathbb{N}$ called the *rank* function, such that $\rho$ satisfies the following two properties:

- If $x, y \in P$ and $x \prec y$, then $\rho(x) < \rho(y)$.
- If $y$ covers $x$ then $\rho(y) = \rho(x) + 1$.

If $x \in P$ and $\rho(x) = i$ then we say that $x$ has rank $i$ and we write $P_i$ for the set of elements of $P$ that have rank $i$.

**Definition 6.2.1.** [Sta90] Let $r$ be a positive integer. A poset $P$ is called $r$-differential if it satisfies the following three conditions:

(S1) $P$ is locally finite, graded and has a unique minimal element 0.

(S2) If $x \neq y$ in $P$ and there exist exactly $k$ elements of $P$ which are covered by both $x$ and $y$, then there are exactly $k$ elements of $P$ which cover both $x$ and $y$.

(S3) If $x \in P$ and $x$ covers exactly $k$ elements of $P$ then $x$ is covered by exactly $k + r$ elements of $P$.

One of the easiest example of an $r$-differntial poset noted by Stanley [Sta90] is the *Young poset* $Y$ As a set $Y$ consists of all *Young tableaux* (see [Sta86]) for all partitions of all nonnegative integers. We say that $\lambda \leq \mu$ if $\lambda$ is contained in $\mu$ i.e., if $\lambda = \lambda_1 + \lambda_2 + \dots$ and $\mu = \mu_1 + \mu_2 + \dots$, where the $\lambda_i$'s and $\mu_i$'s are nonincreasing, then $\lambda_i \geq \mu_i$ for all $i$. Stanley [Sta90] proves that $Y$ is *1-differential*.

A *saturated chain* is a sequence $x_1 \prec x_2 \prec \cdots \prec x_k$ of elements of $P$ such that each $x_{i+1}$ covers $x_i$. The combinatorial properties of an $r$-differential poset $P$ are determined by counting the number $\alpha(0 \to k)$ of saturated chains of the form

$$0 = x_1 \prec x_2 \prec \cdots \prec x_k$$

where 0 is the unique minimal element of $P$. The combinatorial results regarding saturated chains of differential posets in [Sta88b] can be expressed succinctly in terms of generating functions. In particular if $P$ is an $r$-differential poset then

$$\sum_{k\geq 0} \alpha(0 \to k) \frac{x^k}{k!} = \exp\left(rx + \frac{1}{2}rx^2\right),$$

hence there is a rigidity of sorts in this definition.

Stanley expanded the notion of a differential poset in [Sta90] to a larger family called *sequentially differential posets*. Let $\mathbf{r} = \{r_i\}_{i\geq 0}$ be a sequence of non-negative integers. A sequentially $\mathbf{r}$-differential poset $P$ satisfies (S1), (S2) and the following modification of (S3):

(S3)′ If $x \in P_i$ and $x$ covers exactly $k$ elements of $P$ then $x$ is covered by exactly $k + r_i$ elements of $P$.

For example, if $\mathbf{r}$ is the sequence defined by $r_i = r$ for all $i$ then a sequentially $\mathbf{r}$-differential poset is just an $r$-differential poset. Sequentially differential posets retain many of the basic properties of differential posets. The combinatorics of sequentially differential posets are also determined by counting saturated chains but these no longer involve generating functions and thus become much more complicated [Sta90]. For example, if $P$ is a sequentially $\mathbf{r}$-differential poset then there is an exponential generating function for $\alpha(0 \to n)$, namely

$$\sum_{k\geq 0} \alpha(0 \to k) \frac{x^k}{k!} = \sum_{\sigma \in \mathrm{Inv}(k)} \prod_{i \in \mathrm{Wex}(\sigma)} r_{\eta(\sigma,i)}. \tag{6.2.1}$$

Hoffman proved in [Hof03] that the set of rooted trees on a finite number of vertices could be regarded as a sequentially $\mathbf{r}$-differential poset such that $r_i = i + 1$ for all $i \in \mathbb{N}$. We call such a poset a *sequentially increasing differential poset*. We remark that Equation 3.2.3 is an application of Equation 6.2.1. Theorem 4 can thus be generalised as follows:

**Proposition 6.2.2.** If $P$ is a sequentially increasing differential poset then there exists a generating function for $\alpha(0 \to k)$, namely

$$\sum_{k \geq 0} \alpha(0 \to k) \frac{x^k}{k!} = \frac{1}{(1 - x - \frac{x^2}{2})}. \tag{6.2.2}$$

# References


[AB00] Réka Albert and Albert-László Barabási, *Topology of evolving networks: local events and universality*, Physical Review Letters **85** (2000), no. 24, 5234.

[AB02] ______, *Statistical mechanics of complex networks*, Reviews of Modern Physics **74** (2002), no. 1, 47.

[AH00] Lada A Adamic and Bernardo A Huberman, *Power-law distribution of the world wide web*, Science **287** (2000), no. 5461, 2115–2115.

[AJB99] Réka Albert, Hawoong Jeong, and Albert-László Barabási, *Internet: Diameter of the world-wide web*, Nature **401** (1999), no. 6749, 130–131.

[AZHE10] Martin Aigner, Günter M Ziegler, Karl Heinrich Hofmann, and Pál Erdős, *Proofs from the book*, vol. 274, Springer, 2010.

[BA99] Albert-László Barabási and Réka Albert, *Emergence of scaling in random networks*, Science **286** (1999), no. 5439, 509–512.

[BAJ00] Albert-László Barabási, Réka Albert, and Hawoong Jeong, *Scale-free characteristics of random networks: the topology of the world-wide web*, Physica A: Statistical Mechanics and its Applications **281** (2000), no. 1, 69–77.

[BBCR03] Béla Bollobás, Christian Borgs, Jennifer Chayes, and Oliver Riordan, *Directed scale-free graphs*, Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms, Society for Industrial and Applied Mathematics, 2003, pp. 132–139.

[Bil13] Patrick Billingsley, *Convergence of probability measures*, John Wiley & Sons, 2013.

[BK99] David John Broadhurst and Dirk Kreimer, *Renormalization automated by hopf algebra*, Journal of Symbolic Computation **27** (1999), no. 6, 581–600.

[BLL12] Nantel Bergeron, Thomas Lam, and Huilan Li, *Combinatorial hopf algebras and towers of algebras—dimension, quantization and functorality*, Algebras and Representation Theory **15** (2012), no. 4, 675–696.

[BO04] Albert-Laszlo Barabasi and Zoltan N Oltvai, *Network biology: understanding the cell's functional organization*, Nature Reviews Genetics **5** (2004), no. 2, 101–113.

[Bol13] Béla Bollobás, *Modern graph theory*, vol. 184, Springer Science & Business Media, 2013.

[Bro00] Ch Brouder, *Runge–kutta methods and renormalization*, The European Physical Journal C-Particles and Fields **12** (2000), no. 3, 521–534.

[Bro04] Christian Brouder, *Trees, renormalization and differential equations*, BIT Numerical Mathematics **44** (2004), no. 3, 425–438.

[But08] John C Butcher, *Numerical methods for ordinary differential equations*, John Wiley & Sons, 2008.

[C+04] Peter J Cameron et al., *Automorphisms of graphs*, Topics in Algebraic Graph Theory **102** (2004), 137–155.

[CK98] Alain Connes and Dirk Kreimer, *Hopf algebras, renormalization and noncommutative geometry*, Communications in Mathematical Physics **199** (1998), no. 1, 203–242.

[CK99] ______, *Hopf algebras, renormalization and noncommutative geometry*, Quantum field theory: perspective and prospective, Springer, 1999, pp. 59–109.

[CK00] ______, *Renormalization in quantum field theory and the riemann–hilbert problem i: The hopf algebra structure of graphs and the main theorem*, Communications in Mathematical Physics **210** (2000), no. 1, 249–273.

[ER59] Pál Erdős and Alfréd Rényi, *On random graphs I*, Publ. Math. Debrecen **6** (1959), 290–297.

[ET99] Mark Embree and Lloyd N Trefethen, *Growth and decay of random Fibonacci sequences*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 455, The Royal Society, 1999, pp. 2471–2485.

[FK60] Harry Furstenberg and Harry Kesten, *Products of random matrices*, The Annals of Mathematical Statistics **31** (1960), no. 2, 457–469.

[Ful09] Jason Fulman, *Mixing time for a random walk on rooted trees*, The Electronic Journal of Combinatorics **16** (2009), no. 1, R139.

[GPK94] Ronald Lewis Graham, Oren Patashnik, and Donald E Knuth, *Concrete mathematics: A foundation for computer science.*

[Har69] Frank Harary, *Graph theory*, Addison-Wesley Series in Mathematics, Addison Wesley, 1969.

[Hof03] Michael Hoffman, *Combinatorics of rooted trees and hopf algebras*, Transactions of the American Mathematical Society **355** (2003), no. 9, 3795–3811.

[Hol06] Åke J Holmgren, *Using graph models to analyze the vulnerability of electric power networks*, Risk Analysis **26** (2006), no. 4, 955–969.

[Jan05] Svante Janson, *Asymptotic degree distribution in random recursive trees*, Random Structures and Algorithms **26** (2005), no. 1-2, 69–83.

[JBA+06] Marco A Janssen, Orjan Bodin, John M Anderies, Thomas Elmqvist, Henrik Ernstson, RJ McAllister, Per Olsson, and Paul Ryan, *Toward a network perspective of the study of resilience in social-ecological systems.*

[Jor69] Camille Jordan, *Sur les assemblages de lignes.*, Journal für die reine und angewandte Mathematik **70** (1869), 185–190 (fre).

[KN10] Tamás Kalmár-Nagy, *Random walk on a rooted, directed husimi cactus*, Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems–MTNS, vol. 5, 2010.

[KR68] JFC Kingman and AP Robertson, *On a theorem of Lyapunov*, Journal of the London Mathematical Society **1** (1968), no. 1, 347–351.

[KRL00] Paul L Krapivsky, Sidney Redner, and Francois Leyvraz, *Connectivity of growing random networks*, Physical Review Letters **85** (2000), no. 21, 4629.

[LJM+07] Wangzhong Lu, J Janssen, E Milios, Nathalie Japkowicz, and Yongzheng Zhang, *Node similarity in the citation graph*, Knowledge and Information Systems **11** (2007), no. 1, 105–129.

[MA06] Ben D MacArthur and James W Anderson, *Symmetry and self-organization in complex systems*, arXiv preprint cond-mat/0609274 (2006).

[Mah08] Hosam Mahmoud, *Pólya urn models*, CRC press, 2008.

[Mey00] Carl D Meyer, *Matrix analysis and applied linear algebra*, vol. 2, Siam, 2000.

[MSGA08] Ben D MacArthur, Rubén J Sánchez-García, and James W Anderson, *Symmetry in complex networks*, Discrete Applied Mathematics **156** (2008), no. 18, 3525–3531.

[MSOI+02] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon, *Network motifs: simple building blocks of complex networks*, Science **298** (2002), no. 5594, 824–827.

[Ner81] Olle Nerman, *On the convergence of supercritical general (CMJ) branching processes*, Probability Theory and Related Fields **57** (1981), no. 3, 365–395.

[New01] Mark EJ Newman, *The structure of scientific collaboration networks*, Proceedings of the National Academy of Sciences **98** (2001), no. 2, 404–409.

[Nor98] James R Norris, *Markov chains*, no. 2008, Cambridge University Press, 1998.

[Pit95] Jim Pitman, *Exchangeable and partially exchangeable random partitions*, Probability theory and related fields **102** (1995), no. 2, 145–158.

[Rio12] John Riordan, *Introduction to combinatorial analysis*, Courier Corporation, 2012.

[RT08] Anna Rudas and Bálint Tóth, *Random tree growth with branching processes—a survey*, Handbook of large-scale random networks, Springer, 2008, pp. 171–202.

[RTV07] Anna Rudas, Bálint Tóth, and Benedek Valkó, *Random trees and general branching processes*, Random Structures & Algorithms **31** (2007), no. 2, 186–202.

[SB13] Josef Stoer and Roland Bulirsch, *Introduction to numerical analysis*, vol. 12, Springer Science & Business Media, 2013.

[SK01] Clément Sire and Paul L Krapivsky, *Random Fibonacci sequences*, Journal of Physics A: Mathematical and General **34** (2001), no. 42, 9065.

[Sta72] Richard P Stanley, *Ordered structures and partitions*, vol. 119, American Mathematical Soc., 1972.

[Sta86] ———, *What is enumerative combinatorics?*, Springer, 1986.

[Sta88a] ———, *Differential posets*, Journal of the American Mathematical Society **1** (1988), no. 4, 919–961.

[Sta88b] ———, *Differential posets*, Journal of the American Mathematical Society **1** (1988), no. 4, 919–961.

[Sta90] ———, *Variations on differential posets*, Institute for Mathematics and Its Applications **19** (1990), 145.

[Vau08] Michael T Vaughn, *Introduction to mathematical physics*, John Wiley & Sons, 2008.

[Vis00] Divakar Viswanath, *Random Fibonacci sequences and the number 1.13198824...*, Mathematics of Computation of the American Mathematical Society **69** (2000), no. 231, 1131–1155.

[Wol16] Wolfram Research, Inc., *Mathematica 10.4*, 2016.

[WS98] Duncan J Watts and Steven H Strogatz, *Collective dynamics of 'small-world'networks*, Nature **393** (1998), no. 6684, 440–442.

[Yu12] Le Yu, *Automorphisms of random trees*, Ph.D. thesis, Drexel University, 2012.

# Symbol Index


$\mathbf{a} = (a_1, a_2, \ldots, a_n)$: cycle type, 45
$A \wr B$: wreath product, 15
$A_i^k$, 62
$\alpha(0 \to k)$: number of saturated chains, 72
$\mathrm{anc}(x)$: ancestor function, 28
$\mathrm{Aut}^s(t)$, 43
$\mathrm{Aut}(t)$: automorphism group, iii
$\mathrm{Aut}^{s,k}(t)$, 44

$B^+(t)$, 6
$B^-(t)$, 14
$\beta(G)$, 29
$b_k$: expected order of the automorphism group of a random recursive tree, 11

$\chi(\tau)$, 37
$\mathrm{Ch}_v$: children of $v$, 7
$C_i^s(\sigma)$, 62
$c_{n,k}$, 52
$C^s(\sigma)$, 61
$\mathcal{C}(t)$: non-elementary subgroup, 9

$d_k(\{t_i\}_{i=1}^n)$: number of occurrences of $k$ in $\mathbf{v}(\{t_i\}_{i=1}^n)$, 27
$d(v, w)$: graph distance, 7

$E(G)$: edge set, 5
$\mathrm{Ei}(x)$: exponential integral, 57
$e_k$, 68
$\mathcal{E}(t)$: elementary subgroup, 9
$\eta(\sigma, i)$, 68

$F(x)$, 69
$f$: rooted forest, 5
$f \cong f'$: rooted automorphism, 6
$F_k(x)$, 54
$F'(x) := \sum_{n\geq 0} \frac{a_{n+1}}{n!} x^n$: formal derivative, 54

$G = (r, V, E)$: rooted graph, 5
$G = (r, V, E)$: rooted graph
$\bullet$: rooted graph on one vertex, 6
$\mathcal{G}$: rooted ordered trees, 18
$g(\{t_i\}_{i=1}^n)$, 27
$\gamma_k$, 59
$G_k(x)$, 54
$\mathcal{G}_n$: rooted ordered trees on $n$ vertices, 18
$\tilde{g}_n(i)$, 27
$G_{\downarrow x}$: induced rooted ordered subtree rooted at $x$, 23

$\tilde{h}(s_i)$, 28
$h(s)$, 27
$H_i$: holding times, 22
$H_k(x)$, 53

$\mathrm{Inv}(k)$: involutions of $[k]$, 68
$\iota(x_i)$, 61
$\widehat{I}^s(G)$, 39
$\widetilde{I}^s(t)$, 44
$I^{s,k}(t)$, 44

$\tilde{j}(\sigma)$, 43
$j(\sigma)$, 45

$K(t)$: tree multiplicity, 9

$l_n(v)$: labelling, 26
$\mathcal{L}_1$, 43
$\mathcal{L}_2$, 43
$\lambda_{k,i}$, 56
$\Lambda(t)$: root permutation function, 6
$\lambda_w^*$, 22
$\underline{\lambda}_w := \inf\{\lambda > 0 : \widehat{\rho}_w(\lambda) < \infty\}$, 22

$\mathscr{m}$: Malthusian parameter, 36
$m_\tau^\phi$, 37

$\mathbb{N}$: positive integers, 1
$\mathcal{N} := \bigsqcup_{n=0}^\infty \mathbb{N}^n$, 17
$[n] := \{1, 2, \ldots, n\}$, 1
$\mathbb{N}^0 := \{\emptyset\}$, 17
$\mathbb{N}_0$: non-negative integers, 1
$\nu = \mathbb{E}(\xi)$: intensity measure, 36
$\nu_{\mathscr{m}}(\tau)$, 36
$\nu(\tau)$: reproduction function, 36

$|G|$: order, 5
$\mathrm{out}(v, t)$: outdegree of vertex $v$ in graph rooted tree $t$, 7

$P$: poset, 71
$\phi_n(\{t_i\}_{i=1}^n)$, 9
$\pi_w$, 23
$p_k(x)$: recursively exponential polynomial, 55
$\psi(\tau)$, 37
$p_w$, 23

$q_k(x)$: difference polynomials, 55

$\mathcal{R}$: rooted trees, 6
$r(G)$: root, 5
$\mathbb{R}^>$: positive real numbers, 1
$\mathbb{R}^\geq$: non-negative real numbers, 1
$\rho(x)$: rank function, 71
$\rho_{F,k}$, 54
$\rho_{G,k}$, 54
$\rho_{H,k}$, 53
$\rho_w$: density of point process, 22
$\widehat{\rho}_w$: Laplace transform of $\rho_w$, 22
$\mathcal{R}_n$: rooted trees on $n$ vertices, 6
$\mathcal{R}_n^{s,k}$, 44
$\mathcal{R}^{s,k}$, 44

$S(G)$: historical orderings of rooted ordered tree $G$, 19
$\mathcal{SG}_n$: set of historical orderings, 19
$[\sigma]$: conjugacy class representative of $\sigma$, 50
$\Sigma_n$: set of consistent random permutations, 45
$S_n$: symmetric group, 1
$\widetilde{S}_n$, 50
$S_n^{\mathbf{a}}$: permutations with cycle type $\mathbf{a}$, 47
$S_n^{a_1=k}$, 49
$S_n^i$, 62
$S_n^\mu$, 61
sup: supremum, 37
$\mathrm{supp}(\sigma)$: support, 15

$T(\{\sigma_i\}_{i=1}^n)$, 46
$\widehat{T}(\sigma)$, 43
$\widetilde{T}(\{\sigma_i\}_{i=1}^n)$, 46
$t!$: tree factorial, 32
$t$: rooted tree, iii
$G = (r, V, E, <, l)$: ordered, 18
$\widehat{T}|_C$, 61
$\theta(\tau)$, 38
$\{t_i\}_{i=1}^n$: $n$-sapling, 9
$\mathcal{T}_{w,n}$: $n$-saplings, 9
$\{t_n\}_{n\in\mathbb{N}}$: attachment tree, iii
$\widehat{T}(\sigma)$, 47
$t \lhd t'$, 7
$t_{\downarrow v}$: progeny of $t$ at $v$, 14
$\mathcal{T}_w$: attachment trees, 8

$\mathcal{V}$: Viswanath's constant, iii
$V(G)$: vertex set, 5
$\mathbf{v}(\{t_i\}_{i=1}^n) = (v_2, v_3, \ldots, v_n)$: vertex sequence, 27

$\mathbf{v}'(\{t_i\}_{i=1}^n) = (y_1, y_2, \ldots, y_{n-1})$: lower vertex sequence, 46

$\mathcal{W}$, 10
$\widetilde{w}(n)$, 8
$w(n)$: weight function, 7
$\mathrm{Wex}(\sigma)$: weak excedances of $\sigma$ , 68
$w(G, s, i)$ weight of vertex $s_i$, 20
$W_w(G)$: total weight of rooted ordered tree $G$, 20
$W_w(G, s, i)$: historical sequence of total weights , 20

$X_C$, 61
$\xi_w := (\xi_{w,1}, \xi_{w,2}, \ldots)$: point process, 22
$\xi_w(\tau)$: $\xi$-measure of interval $[0, \tau]$, 22
$X_n^{i,k}$, 49
$\widehat{X}_n^{i,k}$, 62
$X_n \xrightarrow{\text{a.s.}} X$: almost sure convergence, 1
$X_w(\tau)$: Markovian pure birth process, 22
$x \prec y$: partial ordering, 71

$\mathcal{Y}_w^d$, 21
$\mathcal{Y}_w(\tau)$, 21
$\mathcal{Y}_{\downarrow x}(\tau)$: induced subtree of $\mathcal{Y}(\tau)$ rooted at $x$, 23

$\zeta(f) = |\mathrm{Aut}(f)|$: order of automorphism group, 6
$\zeta^{s,k}(t)$, 44
$\zeta^s(t) = |\mathrm{Aut}^s(t)|$, 43
$Z_n(t_1, t_2, \ldots, t_n)$: cycle indicator polynomial, 50
$Z_n^{[\sigma]}$, 50
$\{Z_\tau^\phi\}_{\tau \in \mathbb{R}_\geq}$: general branching process with characteristic $\phi$, 37
$z^{\mathrm{type}(\sigma)}$, 50